Theories of Concepts and Ethics

by

John Jung Park

Department of Philosophy
Duke University

Date:_____

Approved:

_____
David Wong, Supervisor

_____
Karen Neander

_____
Wayne Norman

_____
Daniel Weiskopf

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Philosophy in the Graduate School
of Duke University

2013

ABSTRACT

Theories of Concepts and Ethics

by

John Jung Park

Department of Philosophy
Duke University

Date:_____

Approved:

_____
David Wong, Supervisor

_____
Karen Neander

_____
Wayne Norman

_____
Daniel Weiskopf

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy in the Graduate School of Duke University

2013

Abstract

Psychological concepts are mental representations that represent, refer to, or are about properties and categories. In the philosophy of mind/cognitive science, there are various theories of what kinds of knowledge, or information carrying mental states, constitute our concepts. In other words, such theories provide views on what concepts are. The knowledge stored in concepts is thought to be used in the higher cognitive competences such as in categorization, induction, deduction, and analogical reasoning when we think or reason about the extension of the concept. While most concept theories have primarily focused on concrete concepts such as CHAIR, TABLE and DOG, I take such modern theories and apply them to abstract moral concepts such as VIRTUE, RIGHT ACTION, and JUSTICE. I argue for a new overall theory of moral concepts that combines and includes four theories of concepts. This dissertation addresses the question of what the nature of moral concepts is and what further implications this may have in ethics. I contend that our moral concepts may be constituted by prototype, exemplar, theory, and/or emotion-based kinds of knowledge. This tetrad view differs from, for example, certain Humean-based theories that contend that our moral concepts are only constituted by emotions and desires. Finally, I draw further philosophical implications my conclusion may have for applied ethics, normative ethical theory, political philosophy, and meta-ethics.

Examining the nature of concepts is an area in the philosophy of mind/language and cognitive science, but it dates back to at least Early Modern Philosophy. Descartes,

Locke, Berkeley, and Hume called concepts *ideas* that are the building blocks of thought and the human understanding. Such ideas are mental representations that generally refer to properties, categories, and things in the world. Furthermore, they are largely causally responsible for how we think and reason about the extension of our ideas. For example, for Hume, concepts are ideas that are derived from impressions or immediate sense data and are the basis of the human understanding. Hume claims that, "By *ideas* I mean the faint images of these [impressions] in thinking and reasoning… " For Hume, concepts play a functional role in thinking and reasoning. A similar story can be told for Locke in *An Essay Concerning Human Understanding*. Locke writes, "Every man being conscious to himself, that he thinks, and that which his mind is employed about whilst thinking being the *ideas*, that are there…" Locke, for instance, explicitly believes that our idea of PERSONAL IDENTITY, is responsible for how we reason about what the conditions are for being the same person across time. Berkeley claims that our general concept TRIANGLE is constituted by (prototype) knowledge related to such things as having three angles that equal 180 degrees. Such knowledge is acquired based on our personal experiences with particular triangles (that in contemporary terms are mentally represented by exemplars). Berkeley maintains that TRIANGLE itself is constituted by images of particular triangles (exemplars).

My overall dissertation project can be viewed as following in the footsteps of Hume's *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. For Hume, just as the three books that

constitute the *Treatise* synchronically are of the understanding, emotions, and morals, I likewise examine concepts, emotions, then moral concepts. As Hume unifies the books of the *Treatise* by drawing from his first two books to conclude in his final book that moral concepts are constituted by emotions, I likewise engage the contemporary concepts and philosophy of emotions literature in my dissertation to provide a new overall theory on the nature of moral concepts. Moreover, influenced by Newtonian philosophy, Hume attempted to use empirical-based methods in his examination of ethics. In the same general spirit, I in part rely on experimental data to help inform my overall theory of moral concepts.

The first chapter distinguishes psychological concepts from other notions of concepts in philosophy. For example, mental concepts are distinguished from Platonic concepts that are abstract objects. Rather, concepts as understood here are mental states that are responsible for how we cognitively function in the world around us. As we can see from their functional definition, moral concepts are the very heart of moral cognition, but very few moral psychologists have explored the nature of moral concepts even though, unbeknownst to them, they may have run experiments that in some cases can lead to constitution claims on moral concepts. Furthermore, in Lockean terms, concepts are the building blocks or constituents of thought and how the mind is furnished. Concepts are also mental representations that represent things in the world.

Later in this chapter, I discuss the overall views of concrete concepts held by Edouard Machery and Daniel Weiskopf. They argue that there may be various kinds of

knowledge we have that constitute our concrete concepts, where several theories of concepts may be at play. Here, different kinds of knowledge for a concept can be at work in different situations. Discussing Machery and Weiskopf provides a partial outline for the tetrad view of concepts I adopt for moral concepts, where the four different kinds of knowledge that may constitute a moral concept may be used conjointly or separately in cognition depending on the circumstances.

The second chapter examines the contemporary literature concerning the basis of the human understanding and explains the various theories in the literature for concrete concepts. I then take these theories and convert or alter them so they can account for moral features and be applicable to the moral domain and moral concepts. With this conversion, the classical view claims that moral concepts are constituted by bodies of knowledge that represent the necessary and sufficient conditions of a moral category. For example, 'right action' may be constituted by 'an act must maximize the best consequences.' Influenced by Wittgenstein's notion of family resemblance for concepts, the prototype view states that moral concepts are constituted by prototypes or mental representations of the statistically frequent virtues, rules, reasons for action, and features of moral situations found in one's experiences of virtuous individuals and moral acts. For example, my GOOD PERSON concept may be constituted in part by the concepts: JUST, BRAVE, HONEST, and KIND. This cluster of conceptual constituents is not taken to refer to necessary and sufficient conditions. For instance, I may meet a person who is

just, honest, and kind, but not brave. Nevertheless, since she matches most of the features I think good people have, I still classify her as a good person.

The exemplar theory states that concepts are exemplars or bodies of knowledge that refer to particular moral actions or particular exemplary individuals. Like the prototype theory, this view is not understood to claim that exemplars represent necessary and sufficient conditions. For example, my GOOD PERSON concept in part may be made up of the mental representations: MY MOTHER, ABRAHAM LINCOLN, and MARTIN LUTHER KING, JR. The theory-theory claims that concepts are constituted by mental representations of normative ethical theories. Such mental representations occupy defeasible placeholder positions, where such knowledge is consciously or subconsciously taken to be defeasible based on potential further moral learning. Thus, theory knowledge is not considered to be about necessary and sufficient conditions. For instance, my GOOD PERSON concept may be constituted by the particular virtue theory knowledge: ONE WHO HAS AND EXERCISES THE VIRUTES, WHERE THE VIRTUES LEAD ONE TO LIVE A GOOD LIFE.

I also discuss what I call *the emotion theory*, where moral concepts are sentiments and emotions. This view has its roots in Hume and is currently espoused in various forms by philosophers such as Simon Blackburn and Jesse Prinz. For example, Prinz maintains a neo-empiricist view of moral concepts. Neo-empiricist views are modern versions of traditional empiricist theories of concepts held by the likes of Hume, Locke, and Berkeley, where in varying degrees, these three philosophers held that concepts are

pictorial images. Prinz espouses the James/Lange view that emotions are in part perceptions of bodily changes. For instance, fear is the felt perception of trembling and having an elevated heart rate. If moral concepts are emotions, then emotions being perceptions provides a framework for an empiricist or perceptual-based theory of moral concepts.

In the third chapter, I assess the classical, prototype, and exemplar theories of moral concepts. I contend that the classical view is not viable for moral concepts. One argument amongst others against this view is that there are typicality effects found in experimental studies of moral concepts, where some members of a class are considered by subjects to be more typical than other members. However, the classical view implicitly claims that each member should be considered as an equal member since all members equally satisfy the necessary and sufficient conditions for membership. Given that experimental findings do not bear out the predictions of the classical view and that there is no experimental evidence for this theory, classical structure for moral concepts is not psychologically real.

Later, I criticize Mark Johnson, Paul Churchland, Andy Clark, David Wong, Stephen Stich, and Alvin Goldman's claims that individuals actually have prototype structure to their moral concepts. My criticism is made based on the fact that they do not provide the required empirical evidence for these claims that utilize studies particularly on moral concepts. I then justify that many people do have prototype structure by relying on experimental studies specifically on moral concepts such as GOOD PERSON, JUST

PERSON, and MORAL. I also examine successful prototype studies on concepts such as CRIME and LIE; concepts that may not initially appear to be moral concepts because they may not have a normative component to them. By engaging the thin/thick concepts debate with Bernard Williams, John McDowell, Blackburn, and Allan Gibbard, I argue that such concepts are thick moral concepts in that they have both a normative and descriptive component. Thus, they do provide evidence for the prototype theory for moral concepts.

In this chapter I also argue that many individuals do have exemplar knowledge for their moral concepts. Since it has already been proven that many people have moral prototypes and prototypes are a *summary* representation of features, such summary representations must have been derived from representations of particular acts and individuals that were morally noteworthy. Therefore, individuals must have had exemplar knowledge stored in their concepts before forming prototype knowledge.

The fourth chapter attempts to establish that many people at times utilize theory and emotion theory kinds of knowledge. Now, since concepts are functionally defined as playing a causal role in higher acts of cognition, concepts just are those mental states that realize the causal role. This is just like how hearts are functionally defined as pumping blood. Since my heart realizes the causal role of pumping blood, my heart is a heart. This method of functionally identifying concepts to their constituents may be used to prove the viability of the theory and emotion views. I first lay down several qualifications and constraints concerning the use of this functionalist metaphysics of

moral conceptual mental states, where I propound a positive theory on when mental representation influences can be said to actually constitute the concept at hand and when they cannot. I then freshly use this functionalist method in the moral domain to claim the viability of the theory and emotion theories for moral concepts. Several studies show that emotions and representations of normative ethical theories causally influence judgments in the appropriate specified way at various times. Since the moral concepts in the judgment are functionally defined as playing a causal role in categorization *inter alia*, and emotions and representations of ethical theories at various times appropriately realize the causal role, such moral concepts are in part constituted by emotions and theory knowledge, respectively, via the functionalist metaphysics of moral conceptual mental states.

One of the main topics discussed in the fifth chapter is concept combination. Concept combination is concerned with how concepts can combine with each other to form complex thoughts. This chapter addresses several objections from Jerry Fodor that the prototype view cannot account for combination. By relying on my overall tetrad view of moral concepts, where moral concepts can be constituted by different kinds of knowledge that can be used conjointly in various circumstances, I rely on the virtues of the exemplar and theory views. The exemplar and theory views generally can handle concept combination. Hence, insofar as exemplar and theory knowledge simultaneously may also be at work at times alongside prototypes, they can help to address Fodor's contentions against the prototype view.

The final chapter explores the further philosophical implications the previous chapters may have in ethics. As one implication amongst others, I argue that the demise of the classical view shows that many normative ethical theories and debates in normative ethics – such as Hursthouse's virtue ethics, Bentham's utilitarianism, and particular debates in applied ethics and political philosophy – that presuppose that our moral mental representations can have classical structure, must be altered in order to construct a view or conclusion that is psychologically real. I also provide a positive account of how philosophers must change their methodology in deliberating about normative ethics and political philosophy given how we live in a world in which our concepts do not have classical structure. I also criticize the use of conceptual analysis for the motivational judgment internalism/externalism and motivational Humeanism debates.

Contents

Acknowledgements

I would like to sincerely thank my dissertation advisor, David Wong, for looking at countless drafts of my dissertation chapters. Without his tireless effort and support, this dissertation would not have been able to be completed. In alphabetical order, I would also like to sincerely thank my other committee members for their help, efforts, and advice on the dissertation chapters as well: Karen Neander, Wayne Norman, and Daniel Weiskopf. Their insights were invaluable and have greatly helped to strengthen my thesis.

1. Groundwork for Theories of Moral Concepts

While the abstract has provided the overall summary of the dissertation as well as its philosophical import, this first chapter will establish some of the proper groundwork in order to argue for (1) which theories of concepts apply to or are viable for moral concepts and; (2) strong moral concept pluralism. By stating that a theory is viable, I mean that a concept view and its concomitant structure is experimentally supported, and it can be used to successfully describe a structure of a moral concept that a number of individuals actually do have. In this chapter, important distinctions will be made, the definition of concepts will be drawn, projects will be outlined, and the desiderata for theories of concepts will be laid down.

1.1 The Concept/Category & the Concept/Conception Distinctions

An important distinction will be made between concepts and categories. Categories are usually thought to be abstracta that do not exist in our minds; whereas concepts are mental representations of categories.[1] Mental representations are mental states that represent, or, in other words, are *of* or *about* things in the world. For example, there is the class of gold objects with atomic number 79, out there in the world where particular objects like gold nuggets and various gold jewelries fall within this class. Since I have knowledge about gold, in my mind I have the concept GOLD which in

[1] Even categories of things such as dreams and memories are abstract objects, although a particular dream or memory is in the mind.

turn is about the category *gold*. Thus, to mark this distinction between concepts and categories, when specific concepts are discussed they will be marked in capital letters such as the concepts MORAL and IMMORAL. However, when categories are at issue, they will be italicized.

Another important and common distinction made is the concept/conception distinction. One way to understand this distinction is that the purpose for theories of *concepts* is to attempt to determine how such mental representations correctly refer to properties or categories as well as to determine the meaning of concepts.[2] Thus, concept theorists will try to provide a theory of reference and meaning for concepts. As an example of a concept theory, Fodor's informational atomism holds that most lexical concepts, or one word concepts, are primitive atoms that do not have any conceptual structure to them.[3] These symbols have no semantically interpretable component parts and are not decomposable. For example, on this view, LION is not constituted by FOUR-LEGGED, MANE, LARGE, CARNIVORE, and YELLOW. Rather, such information is considered to be merely collateral data. Fodor attaches an informational theory of content or reference to his atomism where one has the concept *C* of the property *P* so

---

[2] As examples of those who explain this particular take on the distinction, see James Higginbotham, "Conceptual Competence," in *Philosophical Issues*, Vol. 9, Concepts (1998), pp. 149-162. Maite Ezcurdia, "The Concept-Conception Distinction," in *Philosophical Issues*, Vol. 9, Concepts (1998), pp. 187-192.

[3] Jerry Fodor, "Information and Representation," *Information, Language, and Cognition*. Ed. P. Hanson, (Oxford: Oxford University Press, 1991). Jerry Fodor, *Concepts: Where Cognitive Science Went Wrong*, (Oxford: Oxford University Press, 1998). Jerry Fodor, *LOT2*, (Oxford: Oxford University Press, 2008).

long as one stands in the proper reliable causal relationship with *P*.[4] In other words, just as the numbers of rings in a tree carry information on the tree's age, a concept carries information about *P* if the concept is under the nomological control of *P*. Here, a concept's reference and meaning is determined by reliable causal relations. For example, the concept LION expresses the real-world property of *being a lion* because lions are the reliable cause of LION-tokenings. Contrary to a Cartesian first-person perspective, one need not believe anything in particular about lions so long as one's tokening of the concept stands in the proper mind-world relation to the property of *being a lion.* Thus, this theory is non-cognitivist or non-epistemic regarding concept possession.[5]

There also are variations on what are understood to be *concept* theories. Concept theories may be thought to be concerned with immutable *correct* concepts that explain

---

[4] For more on the informational view, see Fred Dretske, *Knowledge and the Flow of Information,* (Cambridge, MA: The MIT Press, 1981).

[5] One important problem for informational atomism known as the Disjunction Problem is that erroneous applications of a concept may still be reliable. For example, when hunting in a dense forest I may mistake a man in the distance for a bear and token the concept BEAR. Such a mistake is natural and is often made by others in similar circumstances. Thus, BEAR has two reliable causes and carries information about bears and men. Here, the dilemma for an informational account arises in the question of why BEAR expresses the property of *being a bear* and not the disjunctive property *being a bear-or-man*. Fodor responds to this issue with his asymmetric dependence theory, where the possible ancillary lawful relations a concept may have, such as the relation between BEAR and *being a man*, are asymmetrically dependent on the nomological relation between the concept and the actual property it expresses. Returning to the example, there is a lawful relation between BEAR and the property *of being a man*, but this relation holds only because of the existence of the more fundamental nomological relation between BEAR and the property of *being a bear*, but not vice versa. This asymmetry is Fodor's explanation for why BEAR expresses the property of *being a bear* rather than of *being a man*. If this adequately addresses the Disjunction Problem, then informational atomism may successfully explain intentionality or reference through its mind-world reliable relations.

metaphysical facts such as what actually makes something a chair. Here, concepts are not presupposed to be mental representations and can themselves be abstract objects. Correct concept theories attempt to give the specifications of the metaphysical conditions for something to actually be categorized under the concept. Examples of such concept theories arise from Christopher Peacocke and Georges Rey who believes that concepts are abstract objects and have necessary and sufficient conditions for determining whether an entity may be classified under the concept.[6]

Some theories that are generally understood to fall on the *conception* side of this distinction are the prototype, exemplar, and theory-theory.[7] These, on the other hand, are epistemic theories about how we, for example, judge something to be a chair. Thus, having a conception of a concept is for a subject to be in the epistemic state of having beliefs about the extension of a concept. We may think of conceptions as pertaining to mental representations that, depending on the person, may change over time and may even be incorrect.

Let us turn to an example to better illustrate the distinction between concept views that provide a theory of reference and conception theories. Some small children refuse to call tined utensils "forks" if such utensils are not part of their own household's set of

---

[6] Christopher Peacocke, *A Study of Concepts*, (Cambridge, MA: The MIT Press, 1992). Georges Rey, "Concepts and Stereotypes," *Cognition* 19, 1983.

[7] There are exceptions where some conception theorists such as Susan Carey take their views to also provide a theory of reference. Susan Carey, *The Origin of Concepts*, (Oxford: Oxford University Press), 2009.

silverware.[8] Thus, other family's tined utensils are thought to be pseudo-forks while one's own family's tined utensils are forks. Hence, such children's conception of FORK is in one way or another related to tined utensils that belong only to one's own house. However, though we may think this conception to be wrong, it still may be the case that they possess the concept FORK if we assume something like Fodor's informational atomism, and the children in question stand in the proper reliable mind-world relation to forks. Thus, one can possess the concept of a fork, but one can at the same time have an inaccurate conception of forks. Through further evidence gathering, misapplication, and improvement-making of one's conception, one can form a more appropriate conception.[9]

---

[8] Higginbotham, ibid.

[9] It must be noted that the concept/conception distinction may be controversial. For example, Jesse Prinz and Andy Clark have argued that there is no strict divide between concept theories that provide theories of reference and epistemic conception theories. There may be an overall theory that encompasses both sides and will be more powerful than those views that do not if the overall theory can accomplish this dualistic aim. On the other hand, Fodor adheres to this distinction by arguing that concepts are for representing or having the ability to think about the concept's extension, where any constitutive links between thinking and epistemic capacities such as categorization are severed. For Fodor, in order to possess a concept of *x*, one must have the ability to think about x, where concept possession occurs independently of any epistemic capacities. As discussed, for Fodor, in order to think about *x*, one must stand in the proper causal relation with *x*. Machery can be read as adhering to this distinction as well in that he talks about different goals philosophers and psychologists have. Philosophers have the goal of providing theories of reference for concepts and to explain what makes it the case that we can have propositional attitudes about the object of our attitudes. Meanwhile, the goal for psychologists deals with epistemic capacities and is to characterize the bodies of knowledge that are used by default in the processes underlying the higher cognitive competences (this goal for psychologists will be explained in the next section of this chapter). However, we shall bypass this debate due to the reason that this distinction will be presupposed here mostly as a matter of specifying and limiting the scope of this dissertation. Even if a view should provide a theory of reference as well as account for our epistemic cognitive capacities, this dissertation is strictly limited to the so called conception aspect. Jesse Prinz and Andy Clark, "Putting Concepts to Work: Some Thoughts for the Twenty-first Century," *Mind & Language* 19: (2004), 57-69. Fodor, *Where*

Due to the amount of space the dissertation will take if concept and conception theories are discussed, this essay is primarily focused on conception theories only, and from here forward, we may understand the term "concept" to be about conceptions unless otherwise noted.[10]

1.2 Concepts Defined

Locke understands concepts as being the bare materials or constituents of thought. For example, the principles KILLING IS WRONG and DO NOT STEAL are each constructed of three concepts. Each of the individual concepts is the building block required in order to have these thoughts. However, Machery offers an alternate understanding that is consistent with Locke's and is one which he believes plays a useful role in the practice of psychology. He gathers this definition by putting puzzle pieces together from the works of several psychologists such as Lawrence Barsalou, Karen Solomon, and Douglas Medin.[11] His definition of a concept is:

Cognitive Science. Edouard Machery, *Doing Without Concepts*, (Oxford: Oxford University Press), 2009. Jesse Prinz, "The Return of Concept Empiricism," in *Categorization and Cognitive Science*, ed. H. Cohen and C. Leferbvre, (Elsevier). Jerry Fodor, "Having Concepts: a Brief Refutation of the Twentieth Century," *Mind & Language* 19: (2004), 29-47.

[10] In the philosophical literature, using the term "concept" in place of "conception" may be considered to be an irregular substitution. However, it is a regular substitution for those philosophers working primarily on conceptions.

[11] L. Barsalou, K. Solomon, L. Wu., "Abstraction in Perceptual Symbol Systems," *Philosophical Transactions of the Royal Society of London: Biological Sciences* 358 (2003): 1177-1187. K. Solomon, D. Medin, E. Lynch. "Concepts Do More Than Categorize." *Trends in Cognitive Sciences* 3 (1999), 99-

> A concept of *x* is a body of knowledge about *x* that is stored in long term memory and that is used by default in the processes underlying most, if not all, higher cognitive competences when these processes result in judgments about *x*.

Here, "knowledge" is used by Machery in the psychological sense where it means any contentful[12] or information-bearing mental state that may be used in such higher cognitive competences as categorization and induction. It is not to be confused with the traditional philosophical definition of true justified belief. This psychological understanding of "knowledge" will also be adopted throughout this dissertation unless otherwise specified. Long term memory is memory that can last for days or decades while working memory refers to the relatively small amount of information one can attend to or maintain in a rapidly accessible state at one time.[13] The higher cognitive competences in psychology are controversially distinguished from the lower competences, where both kinds of competences are defined functionally. Examples of lower competences are vision and motor planning, while instances of higher competences are categorization and induction. Machery differentiates them by stating that the lower deal largely with perceptual and motor competences while the higher competences do not usually take perceptual stimuli such as the activation of rods and cones as direct inputs

---

105. Lloyd Komatsu. "Recent views of Conceptual Structure." *Psychological Bulletin* 112 (1992), 500-526.

[12] A contentful mental state is a mental state that has content or represents features of the world.

[13] Barsalou notably diverges from Machery's definition in that he believes concepts are in short term memory. While it is presupposed here that concepts are stored in long term memory, see Machery, ibid. for a rebuttal of Barsalou.

nor do they produce direct motor outputs.[14] Moreover, the higher cognitive competences are to a certain extent under intentional control and their products can be conscious. The higher competences also are slower than the lower competences. By "default," Machery means that concepts preferentially and presumptively are used in most of the higher competences such as deduction, induction, planning, categorization, and analogy-making. In other words, bodies of knowledge stored in a concept of *x* are preferentially used and readily available when thinking or reasoning about *x* rather than being knowledge that is less available and that does not come spontaneously to mind when cognizing about *x*. Since working memory, where concepts are recruited and organized into complete thoughts, is understood to be a limited capacity, only default knowledge about x is stored in the concept of *x*. This has the added benefit of allowing us to generally reason and perform higher cognitive functions in real time without having to always systematically select from all the facts we may know about *x*. As Locke and Machery's definitions of concepts are compatible with each other, we shall understand both views to be appropriate for the definition of a concept.[15]

---

[14] An exception is in the last stages of perception when what is perceived is categorized is a higher competence.

[15] They are consistent since when a concept of x stored in long term memory is used by default in inferring the conclusion that "X is larger than Y," x is a constituent of this conclusion or thought. To note, while Machery acknowledges that both definitions are consistent with each other, he argues that his definition plays a more useful role in psychology. See Machery, ibid.

1.3 The Viability of Theories of Moral Concepts

In order to understand the project of determining which theories are viable for moral concepts, we will first need to understand the Principle of Abstract Concepts. Much empirical psychological work on concepts has studied such theories in light of what are normally thought of as concrete concepts. Much work has focused on three-dimensional household objects, artifacts, and biological species. Less work has been done in relation to abstract concepts broadly construed,[16] and only a handful of studies have examined moral concepts. Importantly, James Hampton, in a unique study, ran tests on eight different abstract concepts, such as BELIEF, SCIENCE, and CRIME, in order to determine whether they had prototype structure similar to the successful results of finding prototype structure in concrete concepts.[17] To note here, the explanation of the various theories of concepts will be conducted in the second chapter, and a full understanding of the content of these views will not be necessary or important in this first chapter. Now, in Hampton's study, the results were a mixed bag where some abstract concepts did show prototype structure, but others did not. For example, SCIENCE and CRIME showed prototype structure while abstract concepts such as A BELIEF and AN INSTINCT, that may

---

[16] Of note, George Lakoff and Rafael Nunez explore the cognitive foundations of mathematical concepts. However, their approach is based on a cognitive metaphor analysis more so than from the vantage point of what are considered the standard theories of concepts such as prototype, exemplar, and theory-theory. George Lakoff and Rafael Nunez, *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics Into Being*, (New York: Basic Books), 2000.

[17] An explanation of Hampton's feature listing task used in this abstract concept study will be explained in chapter 2, section II. James A. Hampton, "An Investigation of the Nature of Abstract Concepts," *Memory & Cognition* Vol. 9 (2), 1981, pp. 149-156.

intuitively be thought to have prototype structure, as a matter of fact do not have such structure. Thus, the upshot from Hampton's experiments is that we cannot safely presuppose that abstract concepts will have the same theoretical concept structure and cognitive processing as those for concrete concepts. As a matter of caution, we cannot draw conclusions about moral concepts solely based on the findings of concrete concepts. Therefore, further work is required in order to ascertain the structure of moral concepts.

A new principle that I derive from Hampton's study, which is a principle he in no way mentions, is what we may call *The Principle of Abstract Concepts* (PAC).

> **The Principle of Abstract Concepts (PAC)**: Any proposed concept structure for an abstract concept AC cannot be based *only* on non-AC or concrete concept empirical findings. They must at least in part be based on experimental data for AC.

PAC is motivated not only based on Hampton's study but also based on the generally understood demise of the classical theory in the literature in regards to non-mathematical and non-logical concepts.[18] Hampton's study shows that we cannot be quick to jump to conclusions that since, for example, it may be thought that KNOWLEDGE or PERSONHOOD does not have classical structure that it therefore must have some specific concept structure that is predominantly found in studies for concrete concepts. For instance, we cannot directly conclude that KNOWLEDGE or PERSONHOOD has a prototype structure. Rather, for *all* abstract concepts, further premises are required to make arguments that a particular abstract concept has such and such a concept structure.

---

[18] A clarification of the demise of the classical view will be discussed in the third chapter.

What are further required by PAC in order to make conclusions about an abstract concept's structure may be experimental studies on the particular type of abstract concept, or they may be such studies coupled with philosophical arguments. As stated, it is also important to note that PAC does not necessarily deny that the empirical findings on concrete concepts can be used to draw conclusions concerning abstract concepts. Such empirical findings may be used to make inferences only in conjunction with specific experimental work that is focused on the type of abstract concept in question. However, PAC does deny that concrete concept empirical work can by itself be sufficient to establish a concept structural conclusion for abstract concepts.

For the most part, when philosophers or psychologists offer or support a theory or theories of concepts, they are usually only concerned with views that pertain to concrete concepts.[19] Based on PAC, these studies do not lead to moral concept structural conclusions. In fact, the handful of explicit moral concepts studies or, in other words, psychology experiments specifically designed to reach concepts-based conclusions has only been for the prototype view, and even some of these tests on their own lack the rigor and systematicity to make a complete case that ethical concepts are prototypes.[20] With the numerous extant theories of concepts out there, there is unquestionably a sizable lacuna in the concepts literature for moral concepts. However, it is my intention to help

---

[19] A Notable exception is Jesse Prinz who offers a neo-empiricist account of moral concepts based on philosophical argument and experiments in moral psychology that are not explicit concept studies. *Furnishing the Mind*, (Cambridge, MA: The MIT Press, 2002) and *The Emotional Construction of Morals*, (Oxford: Oxford University Press, 2008).

[20] An examination of these prototype studies will be conduction in chapter 2.

fill in this ethics gap. Elizabeth Anscombe has stated that there can be no Ethics that is not grounded in a proper Philosophy of Mind.[21] This dissertation attempts to make a significant contribution to this endeavor.

The third and fourth chapters will be devoted to first establishing which theories of concepts are viable for moral concepts. As a qualification, while the extant theories of concepts may not be able to account for all the evidence that may exist for the nature of moral concepts and it may be the case that many new moral concept theories must be constructed due to this fact, the focus of this essay will be limited to the examination of moral concepts in light of the extant concept theories. Given PAC, there generally has not been sufficient support or an adequate examination of which psychological theories of concepts apply to ethical concepts. Many moral philosophers engage in analysis of moral concepts, but few have addressed which possible theories are viable for such concepts in a manner consistent with PAC. Here, one of the main aims of a concept theory is to provide an account of the structure of concepts, where different theories provide different views of a concept's structure.[22] A concept structure explains what constitutes a concept, where a concept generally is thought to be constituted by other more basic concepts and mental representations. Instances of those few who have used such experimental data to examine the nature of ethical concepts are Mark Johnson, Jennifer Frei, and Philip Shaver for prototype structure and Jesse Prinz, using studies that

---

[21] Elizabeth Anscombe, "Modern Moral Philosophy," *Philosophy*, 33: 1958, 1-19.
[22] A list of desiderata for concept theories is given in section 1.5 of this chapter.

are not explicitly concept experiments, for an emotion-based structure, where concepts are constituted by sentiments and emotions.[23] However, to my knowledge, no one has attempted an actual comprehensive examination of our moral concepts by taking into account all the potential theories of concepts. For example, to my knowledge no one has given evidence consistent with PAC for the exemplar view for moral concepts. No thinker has even explained what the theory-theory for concepts will look like in the moral domain, let alone has anyone explicitly stated that such theories are viable for moral concepts. The first aim of this dissertation is to provide this comprehensive analysis of which concept theories are viable for moral concepts by examining each extant psychological theory of concepts in turn. By using 1) the existent moral prototype concept studies; 2) philosophical arguments along with 1) for the viability of the prototype and exemplar theories; and 3) studies in moral psychology on folk moral judgments that are not explicitly designed to draw conclusions in the domain of concepts but do in fact lead to viability concept conclusions for the theory-theory and the newly devised emotion theory; the prototype, exemplar, theory-theory, and emotion views will be posited as the viable theories for moral concepts in a manner that remains true to PAC. Moreover, at times, new analysis will be given for the viable theories of moral concepts as to what the relevant moral rather than concrete concept components of the concept will even look like. As we can see, this dissertation is in part philosophically interesting in

---

[23] More precisely, Johnson argues for a fusion prototype and theory-theory structure for moral concepts. Mark Johnson, *Moral Imagination*, (Chicago: The University of Chicago Press), 1993.

that it will fall at the intersection of concepts and ethics. As Locke, Berkeley, and Hume can be read in varying degrees as believing that our concepts are generally picture-like images, the nature and structure of our concepts too will be explored here. However, the ethics dimension enters the equation as our focus will be on ethical concepts.

Only explicit experimental data for the prototype theory for moral concepts exists. Therefore, in assessing which theories are viable for ethical concepts, initial and additional work is necessary in demonstrating that there is an experimental and, at times, philosophical grounding for all the viable non-prototype theories of ethical concepts, a basic grounding that is readily available to those philosophers working in the concrete concepts domain. As stated, the empirical evidence for specifically moral concepts is rather thin. However, interestingly enough, some studies in empirical moral psychology, while not built specifically for the study of moral concepts, can be used to draw conclusions in the area. The insight that many non-concepts-based studies in moral psychology can be related and used to draw conclusions on moral concepts will be crucial and essential for determining which theories of concepts are viable for moral concepts. Thus, the first aim of the dissertation is to determine which theories of concepts are viable for moral concepts in a manner that is consistent with PAC.

1.4 Moral Concept Pluralism

The second aim of this dissertation, which is related to the first aim, is to establish moral concept pluralism. Moral concept pluralism states that we do not have one

universal structure for all moral concepts that mentally represent the various moral categories, where there is only one concept for a category.[24] Rather, one may have several individuated concepts of a moral category, where concepts are mental representations of categories and where each concept contains a different structure from the others. Here, what structure or concept that functions or plays a causal role in, *inter alia,* categorization, concept combination, and induction depends on the situation. In order to arrive upon moral concept pluralism, one must first establish that there is more than one viable structure for a possible concept(s) of a category that is actually used in various facets of cognition such as categorization, concept combination, and induction. Only after this may one then show that there is a pluralism rather than hybridism between the viable structures. One difference between pluralism and hybridism is that for pluralism, the viable structures each constitute disparate individuated concepts of a category. Thus, if the prototype, exemplar, theory-theory, and emotion views are the viable fundamental structures of VIRTUOUS for an individual, then one has four individuated concepts of VIRTUOUS, where each concept exclusively contains one of the structures. On the other hand, in a hybrid theory, all the fundamental structures are parts of the same concept. Thus, in the example, if VIRTUOUS is a hybrid, then there is only one concept VIRTUOUS that contains four parts that correspond to the four viable structures. A further elaboration of this particular issue of intrapersonal

---

[24] There are other different types of pluralism for concepts that will not be the focus here. See Gualtiero Piccinini and Sam Scott, "Splitting Concepts," *Philosophy of Science* 73: (2006), 390-409.

concept individuation and the difference between pluralism and hybridism will be provided shortly.

To expand on pluralism further, it is not the case that, for example, all ethical concepts such as BAD PERSON and JUST will always have prototype structure for all people at all times and places. Rather, it may be the case that individuals have a different conceptual structure for a particular moral concept as compared to other individuals. Pluralism also maintains that synchronically at a particular time, an individual may have several possible structures or concepts of a category. Moreover, it may also be the case that diachronically across time, individuals may have different conceptual structures or concepts of a category as well. To give examples of pluralism, in our debate of whether a given war is just or not, I may have a prototype of JUST while you may have a theory-theory structure of it. In this case, we both have different conceptual structures of JUST. Furthermore, it could be the case that I have both a prototype and an exemplar conceptual structure or two individuated concepts of JUST, but in our discussion, I happen to be only relying on my prototype concept. Also, when I think of this issue by myself a month from now, I may have theory-theory and emotion view concept structures or two such concepts of JUST in addition to the two individuated prototype and exemplar structured concepts of the class.

Two philosophers who may be read as arguing for concrete concept pluralism are Edouard Machery and Daniel Weiskopf.[25] Machery believes in what he calls the Heterogeneity Hypothesis, which is really a family of views:

1) Synchronically or diachronically, individuals have several concepts of a category that are used by default in the higher cognitive competences.
2) Co-referential but structurally different concepts or bodies of knowledge have few properties in common.
3) Prototype, exemplar, and theories are the fundamental conceptual structures.
4) Concept structures in 3) often are used in distinct cognitive processing.
5) The notion of "concept" is not a natural kind and ought to be eliminated.[26]

Although he does not explicitly call himself a "pluralist," we can read Machery as being a proponent of pluralism in that he holds 1).[27] On the other hand, 2) is provided in order to support 5) in that if structurally different co-referential bodies of knowledge have few properties in common, then it will be difficult for "concept" to be a natural kind. Since our discussion here is of pluralism and not necessarily one of concept eliminativism, we shall ignore 2) and 5). While eliminativism may or may not be a result

---

[25] Edouard Machery, "Concepts are not a Natural Kind," *Philosophy of Science* 72, (2005), 444-67. Edouard Machery, *Doing Without Concepts*, (Oxford: Oxford University Press, 2009. Daniel Weiskopf, "The plurality of concepts," *Synthese* (July 2009). "Atomism, Pluralism, and Conceptual Content," *Philosophy and Phenomenological Research*, 79: (2009), pp. 130-162. See also D. Medin, E. Lynch, and K. Solomon, "Are There Kinds of Concepts?" *Annual Review of Psychology* 51: (2000), 121-147.

[26] Machery, *Doing*, ibid.

[27] Machery understands pluralism differently than as defined here. He takes pluralism to be "the view that a natural kind K divides into several natural kinds, K_1,…, K_n (*Doing Without Concepts*, 240)." It appears that he understands pluralism to maintain that K is a natural kind that should not be eliminated. This runs contrary to his position that the notion of concept ought to be eliminated. Thus, he is not a pluralist in his own sense of the term. However, pluralism as used by Weiskopf and as defined in this dissertation is merely about Machery's 1) with no further statement about whether concept eliminativism does or does not follow from it. Hence, we may read Machery as a pluralist as defined here.

of pluralism, these two issues may be separated, and such a separation will be made now. 3) lists what Machery believes are the viable theories of concepts for concrete concepts. Since this essay is concerned with abstract moral concepts rather than concrete concepts, 3) does not apply here. Nevertheless, as will be shown, the set of fundamental concepts for ethical concepts include the three given by Machery as well as what will be called the emotion view. Since 4) is linked with 3), it too is inapplicable here, although there is agreement on 4). Thus, we may interpret that Machery believes in pluralism since he holds 1), which is that individuals have several concepts for a category, synchronically or diachronically.

Weiskopf holds that pluralism is the view that one may have several concepts of a category that are each composed of different fundamental psychological structures synchronically or diachronically.[28] Weiskopf also reads Machery as a pluralist, but differs from Machery in that he does not believe concept eliminativism follows from pluralism. Weiskopf states that when used in the higher competences, different structures or bodies of knowledge will be at work depending on the situation at hand. In this sense, Weiskopf claims that our overall conceptual system appears to be adaptive. The general outline of his pluralism argument is:

P1: Psychological theories of concepts state that concepts play a causal/explanatory role in cognition.
P2: Concepts are identified with the structure that best fills those causal/explanatory roles.
P3: Several distinct kinds of structure satisfy the causal/explanatory roles.

[28] Weiskopf, ibid.

C: Therefore, concepts are constituted by several distinct kinds of psychological structures.[29]

Moral concept pluralism must be distinguished from a hybrid theory of moral concepts. For, assuming that there are several viable theories of moral concepts, this may provide evidence for a hybrid rather than a pluralistic view. For instance, since there may be numerous viable conceptual structures in the moral domain, a hybrid moral concept such as RIGHT ACTION may be thought to be simultaneously constituted by all those various viable knowledge structures as parts. The philosophers who have done most to outline the distinction between hybrid and pluralistic views of concepts are Machery and Weiskopf.[30]

Machery examines Osherson and Smith's classic hybrid theory in order to draw his distinction. Osherson and Smith believe concepts are composed of a core and identification procedure parts.[31] The core aspect of a concept has classical structure and is composed of a body of knowledge that is about the necessary and sufficient features of members of a class. The identification procedure aspect has prototype structure that is made up of a body of knowledge that represents the statistically frequent features of members of a class. The core part of a concept is at work in concept combination and categorization when there is no time pressure. This use of time pressure in this particular

---

[29] Ibid.

[30] Machery, *Doing*, 63-75. Weiskopf, *Plurality*.

[31] Daniel, Osherson and Edward Smith, "On the adequacy of prototype theory as a theory of concepts," *Cognition* 9: (1981), 35-58. For another classic hybrid theory, see E.Smith, E. Shoben, L. Rips, "Structure and process in semantic memory: A featural model for semantic decisions," *Psychological Review* 81, (1974), 214-241.

hybrid is a characteristic of some but not all hybrid theories. For example, a hybrid theory alternatively may claim that one structure is used for making correct categorizations while another body of knowledge is used to judge how typical a member is within its category as compared to other members. For Osherson and Smith, the identification procedure part is used when we are asked to categorize items as members quickly with little time for thought. Based on an examination of their hybrid view, Machery identifies hybrid theories as claiming to have the following four properties:

1) A concept is divided into several parts.
2) Each part stores a distinct type of knowledge – e.g., knowledge of prototypes, knowledge of exemplars, etc.
3) When one of the parts is used in a higher cognitive competence, then it is necessarily the case that the other parts of the concept may be used in other cognitive competences.
4) The parts of a given concept do not produce inconsistent outcomes such as inconsistent categorization judgments.[32]

Osherson and Smith's hybrid is in line with 1) in that a concept is divided into two parts: The core and identification procedure. With 2), the hybrid has two different types of knowledge associated with the two parts: Knowledge about necessary and sufficient features and summary representations features. Moreover, the parts of a concept are linked in that the core is at work in concept combination and necessarily the identification procedure may be used in fast categorization tasks. Thus, 3) is satisfied. Machery motivates 3) based on the fact that traditionally, hybrid theories claim that 3) is the case for hybrid concepts. However, Machery states that while the Heterogeneity

[32] Machery, ibid.

Hypothesis (or more precisely, the first tenet of the Hypothesis) allows for the possibility for different bodies of knowledge of *x* to be used in different competences, such a linkage is contingent rather than necessary, and it may not be the case that, for example, a prototype of *x* may be used in a competence when the classical structure of *x* has been used in some other competence. For 3), we may understand hybrid theorists to be using an empirical type of necessity; a necessity where something could have been otherwise, but it did not or was not the case. Finally 4) is met due to the fact that the two parts are coordinated and at times do not produce inconsistent categorization judgments. In this arrangement, the core may at times override the identification procedure to produce a single categorization output, where such coordination may not allow for two opposing final outcomes from the two different parts of the hybrid. 4) is satisfied because the core is the criterion of correctness for a concept and at times may supersede the identification procedure. While the identification procedure is considered to be a reliable yet defeasible body of knowledge, the criterion of correctness is designated to be used when we must be sure to attempt to categorize correctly. Now, a part is the "criterion of correctness" only in the sense that as a psychological matter, it may override the other parts since it is considered to be the means to categorize and reason correctly. However, it still is possible that the part that is the criterion of correctness may contain misconceptions. While there are different types of hybrid theories that are each set up in different ways,[33]

---

[33] For other hybrid theories see: L. Rips, E. Shoben, and E. Smith, "Semantic distance and the verification of semantic relations," *Journal of Verbal Learning and Verbal Behavior* 12: (1973), p. 1-20.

having a criterion of correctness that may be overridingly used when the potential for inconsistencies among the possible bodies of knowledge arises is a general feature of traditional hybrid theories.[34] It is important to keep in mind for Machery that on a hybrid view the possibility for conflicting or inconsistent judgments might be eliminated, where the criterion of correctness itself is a consistent body of knowledge. Meanwhile, concept pluralism denies that there is a criterion of correctness for the entire particular domain of concepts in question. All in all, we can see that Osherson and Smith's hybrid view satisfies the four properties or claims that characterize hybrid theories, where whether moral concepts are hybrids or not will be discussed in the fifth chapter.

As a clarification and further refinement of hybridism, it appears that even if we have a criterion of correctness and have the competence to resolve contradictions but our performance is such that we generally do make judgments consistent with the criterion of correctness but at times make contradictory judgments because of mental error, a hybridist will still claim that the conceptual organization of our minds is that of a

---

E. Smith, E. Shoben, and L. Rips, "Structure and process in semantic memory: a featural model for semantic decisions," *Pscyhological Review* 81: (1974), 214-241. S. Armstrong, L. Gleitman, and H. Gleitman, "What Some Concepts Might Not Be," *Cognition*, 13, 1983. F. Keil, W. smith, D. Simons, and D. Levin, "Two dogmas of conceptual empiricism: Implications for hybrid models of the structure of knowledge," *Cognition* 65: (1998), p. 103-135.

[34] To note, some views call themselves hybrids, but do not posit a criterion of correctness. When hybrids are discussed in this dissertation, they will refer only to those certain theories that do put forth a criterion of correctness. For an example of a so-called hybrid that does not have a criterion of correctness, see F. Keil, W. Smith, D. Simons, D. Levin, "Two dogmas of conceptual empiricism: implications for hybrid models of the structure of knowledge," *Cognition* 65: (1998), pp. 103-135.

hybridism in that we still really do have a criterion of correctness that generally leads us to make the corresponding judgments in a variety of circumstances even though at times there may be performance errors due to some kind of mental mistake. We may understand a competent person in terms of hybridism to be one who really does have a criterion of correctness such that this person's performance of making actual judgments is generally consistent with it, but at times this person may make contradictory judgments due to some kind of mental error. In other words, if one has a criterion of correctness that is competently mastered, then one has the ability to make a certain pattern of judgments under a variety of circumstances that is in line with and is what we should expect from having a criterion of correctness, and one has available plausible explanations of mental error when conflicting judgments are made; explanations that are consistent with mastery of making judgments that are in league with the criterion of correctness under normal circumstances. From here forward, we shall understand the hybridist's claim against conflicting judgments to be along these lines.[35] If one has competence in having and applying a criterion of correctness, then the various types of structures in question may be parts of the same individual concept so long as 3) is satisfied as well. However, if there is no such competence in a given concept realm, then one will have a pluralism and several individuated concepts of a category corresponding to the number of viable structures.

---

[35] The nature of how contradictory judgments may arise will be discussed in chapter four.

Now, mental error may be due to any number of factors such as forgetfulness, mental fatigue, drugs, or perhaps certain reasoning biases of the mind such as the irrational primacy effect in which subjects give more weight or importance to information at the beginning of a series of data even though the order in which the information is given is not important.

Also, there may be a possible intermediate view between pluralism and hybridism. This position may claim that in some sub-domains of a concept realm an agent does have a criterion of correctness that is competently mastered by the user, but in other related conceptual sub-domains the agent does not have such a criterion. For example, in the realm of concrete concepts, we may have such a criterion for natural kind concepts, but we may not have it for artifact kind representations. For ethical concepts, this middle view may state that there is such a criterion for thin concepts such as RIGHT ACTION and WRONG ACTION, but there is no such criterion for thick concepts such as HONEST and PUSILLANIMOUS.[36] This position differs from hybridism in that there is no criterion of correctness that is competently mastered for an agent in at least one subdomain, which is a position that hybridism flatly denies. However, the issue becomes more complicated when comparing it to pluralism. For, within a domain, pluralism merely claims that inconsistencies that are not due to any type of performance errors but rather to a lack of a criterion of correctness are possible, and it may be understood as not

---

[36] As will be further elaborated upon in the third chapter, thin concepts are thought to contain only a normative component while thick concepts contain both a normative and descriptive component.

necessarily denying that there may be a criterion of correctness that is competently mastered in one subdomain but not in another. It may be seen as being consistent with pluralism in that inconsistencies that are not due to any type of mental error are possible on this view in at least one subfield of ethical concepts because there is no criterion of correctness that is competently mastered in this subfield. However, there may be such a criterion in another subfield. Therefore, we shall describe *strong* pluralism as the view that there is no criterion of correctness that is competently mastered at all in a given concept realm. Meanwhile, *weak* pluralism will be this intermediate view in which there may be such a criterion in some subdomains of a given field of concepts, but there still is no such criterion in other subdomains.

Here, we can even further fill up the logical space of possibilities by allowing for the option that hybridism or even weak pluralism may allow for different bodies of knowledge to be the criterion of correctness based on the subdomain of concepts. For example, I may hold a hybrid view in which my prototypes are the dominant representations for natural kind concepts in physics and chemistry, theories are overriding for natural kind concepts in biology, and exemplars are the criterion of correctness for artifact kind concepts. Likewise, if I am a weak pluralist, I may have the same bodies of knowledge play the dominant roles in their respective subdomains as in the above example, except for natural kind concepts in chemistry, where I maintain that there is no criterion of correctness. For this possibility of having several different bodies of knowledge rather than one play the overriding role within subdomains of a general

domain of concepts, we shall entitle the relevant views as *multi-hybridism* and *multi-weak pluralism*.

Now, Machery understands 3) and 4) as part of a theory of intrapersonal concept individuation in that they are individually necessary and jointly sufficient conditions for two co-referential bodies of knowledge to be parts of the same concept as opposed to composing two distinct concepts. In other words, if 3) and 4) are met, then the two types of structure in question are parts of the same individual concept. However, if 3) or 4) are not satisfied, then, for example, Osherson and Smith's hybrid view is really a pluralism view where one will have two individuated concepts of a category, where the first concept will have classical structure and the second concept will have prototype structure. Thus, it will not be the case that the core and identification procedure will hybridize and be parts of the same concept.

Similar to Machery's 1) and 2), Weiskopf states that a hybrid theory of concepts claims that there are several parts to a concept, where each part contains a body of knowledge that has different types of information and processing from the other parts. For hybrids, there is only one concept of a category. He writes, "[H]ybrids are theories on which concepts are identified with *single* representations that possess two or more distinct components that have significantly different characteristics…"[37] Weiskopf also refers to Osherson and Smith's theory to account for his understanding of hybridism. Weiskopf does not recognize Machery's 3), but he does mention along the lines of 4) that

---

[37] Weiskopf, "Plurality," 168.

hybrid theories usually posit that one part of the concept is taken to be the criterion of correctness. While it cannot be conclusively stated that Weiskopf adopts Machery's theory of concept individuation since Weiskopf does not discuss 3) at all nor 4) in regards to individuation, Weiskopf can be read as accepting the conclusion that hybrid theories hold that there is only one concept for a category in which the single concept has several disparate parts that each have different types of structure and information. Meanwhile, a pluralistic view contends that each different viable type of structure constitutes a disparate individuated concept of the category. Thus, like Machery, for pluralism one may have several concepts of a category that each has different types of knowledge.

Based on our hybrid versus pluralism discussion of Machery and Weiskopf, we will adopt Machery's criteria for distinguishing between hybrid and pluralistic views. Moreover, his theory of concept individuation will also be endorsed. Their jointly held view that hybrids posit only one super-structured concept of a category while pluralism, as defined here, contends that there are numerous concepts of a category will also be maintained. In summary, we see that Machery and Weiskopf can be viewed as holding a pluralistic account, where synchronically and diachronically one may have several individuated concepts of a category and where different concepts of a category may be used in the higher cognitive competences depending on the situation. It is this understanding of concrete concept pluralism that will be adopted here for moral concept pluralism.

Both Machery and Weiskopf argue that many to most thinkers working in the concrete concepts field believe that all concepts have a single uniform type of structure. However, it is important to keep in mind that our focus is on the realm of abstract ethical concepts, not concrete concepts. Given that our domain of inquiry is ethics-related, some philosophers and psychologists who have studied moral concepts have held a prototype view. Several proponents are Linda Coleman, Paul Kay, James Hampton, Mark Johnson, and Paul Churchland.[38] Given PAC, since they have not considered the further requisite evidence for other theories of concepts, there are no alternative conceptual structures for such thinkers from which a pluralism may even be recommended. It may be the case that such thinkers believe there are several viable moral concepts structures beyond the prototype view and that there is a pluralism that exists between the fundamental moral structures. This may be the case even though they have not explicitly stated what the other viable structures are or that there is even a moral concept pluralism. Furthermore, this may be the case even though they have not provided arguments as to why there is a pluralism rather than hybridization between the structures. Yet, given PAC, since they have not considered the further requisite evidence for other structures, the questions still

---

[38] Paul Churchland, *A Neurocomputational Perspective*, (Cambridge, MA: The MIT Press), 1989. James Hampton, "An Investigation of the Nature of Abstract Concepts," *Memory and Cognition,* Vol. 9 (2), 1981. Linda Coleman and Paul Kay, "Prototype Semantics: The English Word Lie," *Language* 57:1. (March 1981), 26-44. Abigail Strichartz and Roger Burton, "Lies and Truth: A Study of the Development of the Concept," *Child Development*, 61:1 (Feb., 1990), 211-20. Johnson, ibid. William Casebeer's emphasis on the mind in *Natural Ethical Facts* is on connectionism and moral judgment. He does not discuss the concepts literature in any detail.

remain as to what the other viable structures are as well as to what the argument is for moral concept pluralism. Contrastively, such evidence and arguments will be provided here.

While Jesse Prinz may be read as holding two different theories of moral concepts – a pluralism and a monism – in reconciliation, it will be argued that Prinz actually holds a pluralistic view, where our moral concepts are either composed of exemplars or of sentiments and emotions. However, Prinz does not provide evidence consistent with PAC for the exemplar theory and his case that moral concepts are constituted by sentiments and emotions will be criticized in chapter four. Moreover, Prinz does not give an argument for pluralism over hybridism. Nevertheless, the appropriate arguments for the exemplar and emotion-constitution views that are true to PAC will be given here. Also, the contention for pluralism over hybridism will be provided.

Some philosophers can be read as having adopted a moral concept pluralism such as David Wong, Andy Clark, Alvin Goldman, and Stephen Stich.[39] They may be read as stating that it is plausible that there is a pluralism between prototype and exemplar theory. While I agree with this basic position, this essay provides a justification for this conclusion in line with PAC that is absent from their immediate discussions as well as

---

[39] David B. Wong, *Natural Moralities*, (Oxford: Oxford University Press, 2006), pp. 69-71. A moral concept pluralism may implicitly be given by Andy Clark, "Connectionism, Moral Cognition, and Collaborative Problem Solving," *Mind and Morals*, ed by L. May, M. Friedman, and A. Clark, (Cambridge, MA: The MIT Press, 1996), p.114. Alvin Goldman, "Ethics and Cognitive Science," *Ethics* Vol. 103 (Jan 1993), 337-360. Stephen Stich, "Moral Philosophy and Mental Representation," *The Origin of Values*. Ed. by Hechter, M., Nadel, L., and Michod, R. (New York: Aldine de Gruyer, 1993), pp. 215-228.

augments the purview towards which their pluralism applies by adding the theory-theory and the emotion view as viable moral concept theories. Now, they may not necessarily exclude the possibility that there are more viable moral structures, but this dissertation will give the evidence consistent with PAC for such structures as well as the actual contention for pluralism that is not contained in their writings. Thus, the evidence consistent with PAC for the fundamental moral concepts structures and the actual arguments for strong pluralism rather than hybridism will be provided here. The actual establishment of a strong moral concept pluralism in Meta-ethics, Philosophy of Mind, Philosophy of Cognitive Science, and Philosophy of Language will uniquely be attempted in this dissertation.

1.5 The Desiderata for Theories of Moral Concepts

Since we have a definition of "concept," the desiderata on a theory of concepts will be explored. These are the explanatory roles any promising theory of concepts must to a certain extent satisfy. In the third through fifth chapters, it will be argued that my pluralistic view of moral concepts best accounts for these criteria. Now, since pluralism will be contended, where there are numerous fundamental theories, each individual theory need not wholly satisfy each of the criteria, since other viable theories may account for its deficiencies. Stephen Laurence, Eric Margolis, Fodor, Prinz, and Machery

have done much to clarify what these criteria for theories of concepts must be.[40] A look at their relevant work will be beneficial to the reader. As here our purposes are primarily focused on conception rather than reference-based concept views, the list of goals provided will be tailored to conception theory purposes. Thus, one may expect differences and alterations in the list of criteria compared to some of the above authors, whose lists, for some thinkers, conflate the desiderata of concept and conception theories.[41] It is for this reason that providing a refurbished list of criteria is appropriate in spite of the detailed work done by the specified group of authors. To note, while there may be other criteria relevant to the ethical domain of concepts such as analogical reasoning and planning, due to space concerns, only the standard criteria for theories of concrete concepts that also apply to ethical concepts will be examined.

Before listing the criteria, several common desiderata for theories of concepts will now be discussed as to why they are not included in the given list of criteria. First, a possible criterion of theories of concepts is the scope requirement where, as understood here, one fundamental concept theory, a single hybrid view, or a pluralistic theory of concepts has to be able to account for the different types of concepts that are abstract, concrete, and sensory. For example, a theory of concepts is thought by some to have to be able to encompass a diverse array of concepts such as TRUTH, FEAR, TURTLE,

---

[40] Stephen Laurence and Eric Margolis, "Concepts and Cognitive Science," *Concepts: Core Readings*, Ed E. Margolis and S. Laurence, Cambridge: MA, The MIT Press, 1999, 3-82. Fodor, *Concepts*. Prinz, *Furnishing the Mind.* Machery, *Doing Without Concepts.*

[41] For an explanation of the confusion in the concepts literature of conflating both types of theories, see Machery's first two chapters in: Machery, *Doing*, 7-51.

DOORKNOB, and JUSTICE. This desideratum is considered to be the bane of the traditional empiricism view held by Locke, Berkeley, and Hume. Traditional empiricism's demise is in part brought about due to its failure to account for abstract concepts, such as PUSILLANIMITY, MORAL, WRONG, and VIRTUE, within an image-based theory.

For our purposes we will be examining only abstract ethical concepts. This inquiry is about ethical concepts concerning how we should or should not act inter and intra-personally, and it is also about virtue trait concepts. For now, we may understand it as not being meant to be about conventional normative concepts that are thought to pertain to such actions as not blowing one's nose at the dinner table or always wearing formal clothing when at the clubhouse. As we will see later in the fourth chapter, it is very difficult if not impossible to come up with a principled distinction between the moral and the conventional. However, for now, suffice it to say that this essay concerns itself with normative concepts pertaining to those actions that are usually thought to be moral; although, this in itself is contentious on some issues, and what is considered to be moral may also be likely subject to cultural variance. We also will not be concerned with concepts such as FREE WILL or PERSONAL IDENTITY which are commonly classified as concepts in the domain of metaphysics but may also be thought to fall in the area of moral responsibility. Thus, our area of interest is a highly specified set of abstract concepts and does not pretend to be about all types of concepts concrete, sensory, and abstract. In this respect, the scope requirement does not necessarily apply here.

Based on the concept/conception distinction, conception theories supposedly focus on how human beings epistemically and conceptually categorize the particular as instances of the general. This is not to be confused with the question of how concepts have intentionality or are about and refer to things.[42] Assuming the distinction, concept rather than conception theories tend to be more interested in providing a theory of reference, extension, denotation, or the explanation of in virtue of what concepts may refer to what they refer. The criterion of intentionality generally deals with the issue of concepts and how they refer, where the intentional content of a concept is those things to which a concept refers. For example, recall that for Fodor's concept view, rather than conception view, the concept LION refers to the actual property of *being a lion* based on an informational theory. On the other hand, conception theories are more concerned with how we psychologically or epistemically categorize the more particular actions, entities, or classes into more general or superordinate categories. For example, as will be further explained in the next chapter, the prototype theory of conceptions claims that conceptions are bodies of knowledge that are about the statistically frequent summary representation features of members of a class. Thus, for some individual, LION may be composed of YELLOW, MANE, LARGE, FOUR-LEGGED, and ROARS. However, for most conception theorists, this aspect of the prototype theory is not to be understood as pertaining to some descriptivist theory of reference but rather, as a psychological explanation of how human

---

[42] See Machery, ibid.

beings categorize.[43] Prototype theorists posit such a structure for our concrete conceptions as a way to explain categorization as opposed to intentionality. Thus, prototype theorists are interested in categorization as concerned with how we may categorize *lions* as being subsumed within the larger class of *mammals* by way of matching the features between them. They are also concerned with whether and how test participants categorize particular tokens of large four-legged and yellow predators as falling within the category of *lion* by way of feature matching as well. Thus, conception theories need not be concerned with whether and how LION properly refers to *lion*, as this is an issue for concept rather than conception views. The criterion of categorization may be thought to be similar to but is not to be confused with the criterion of intentionality. Although we will assume such things as that our conceptions do refer or purportedly refer, since our focus is on conceptions rather than concepts, intentionality or providing a theory of reference will not be a desideratum of conception theories, or not at any rate one which will concern us.[44]

The criterion of publicity may be thought to be a desideratum of concepts.[45] This is where concepts must be sharable between human beings in terms of their intentional

---

[43] Even if a conception theorist such as Susan Carey holds that the features of a conception determine reference, this intentionality aspect may in principle be separated from the issue of categorization. Susan Carey, *The Origin of Concepts*, (Oxford: Oxford University Press), 2009.
[44] Machery especially emphasizes this point. Machery, *Doing*, 7-51.
[45] Jerry Fodor, *Concepts: Where Cognitive Science Went Wrong*, (Oxford: Oxford University Press, 1998). Christopher Peacocke, *A Study of Concepts*, (Cambridge, MA: The MIT Press, 1992). Georges Rey, "Concepts and Stereotypes," *Cognition* 19, 1983. Jesse Prinz, *Furnishing the Mind*, Cambridge, MA: The MIT Press, 2002.

content and possibly in their sense or cognitive content as well.[46] It is usually thought that concepts play a role in linguistic communication and that people assume that others share the same concepts with one another in order that they can understand each other and communication is possible. Thus, publicity asks theories to explain how concepts are sharable in light of their intentional content and possibly cognitive content. For example, if we are having a discussion about whether torturing terrorist suspects in order to gather vital information is right or wrong, it may be assumed that we share the same concepts MORALLY RIGHT and MORALLY WRONG. If our moral concepts are not the same as each other's and are not sharable in light of intentional content and possibly cognitive content, then we have different concepts and are not communicating with each other.

What may be seen as a precondition for publicity is the issue of interpersonal concept individuation, which is concerned with what aspects make a concept the particular concept it is such that it may be said to differ from other contrast concepts, which are concepts that are similar yet different from the concept in question. There is within-individual individuation and between-individual individuation.[47] The first kind is what we have already discussed in Machery's criteria 3) and 4) for a hybrid versus a

---

[46] "Cognitive content" will be explained and clarified further below. In order to have publicity, two concepts must share the same extension in order that two individuals are referentially talking about the same thing. However, having differences in conceptual components does not necessarily mean that two co-referential concepts are different concepts and that there can be no publicity. For example, one may contend that concepts are solely individuated by reference or that if there is enough similarity in the two conceptual components, then both individuals can be said to still share the same co-referential concepts, and there can be publicity.

[47] Machery, *Doing*, 14.

pluralistic theory. Intrapersonal or within-individual individuation is concerned with whether two co-referential yet perhaps different bodies of knowledge in the same person compose two different concepts. Between-individual individuation generally deals with whether two co-referential yet perhaps different bodies of knowledge possessed by two individuals are the same concept.[48] Here, it is assumed that two concepts can co-refer despite the possibility that these concepts may be different concepts. Now, we can see how related this latter individuation is to that of publicity in that if your and my concepts of RIGHT ACTION may be individuated from each other, then we both relevantly have different concepts. If we have different moral concepts, then when we discuss whether torturing terrorists to attain vital information is right or not, then it may be the case that we lack publicity.[49] For instance, both of our concepts of RIGHT ACTION may refer to the same moral category but you describe this concept as being mandated from Zeus while I describe it as being the product of human reason. Depending on one's theory of between-individual individuation and assuming that none of us have made any mental errors in our concepts RIGHT ACTION, in this circumstance it may be the case that we have different concepts. If this is so, then we may lack publicity and communication. For our purposes, we will not pursue publicity and between-individual individuation. As a concern for limiting the breadth of this essay, the list of criteria for moral concept theories will

---

[48] For between-individual individuation, if the concepts at hand are not co-referential, then they are different concepts since the concepts are literally about different things.

[49] Although our concepts will importantly share reference, if our concepts can have enough differences from each other in terms of sense or cognitive content, then it may be the case that our concepts are still different concepts. Thus, we may lack publicity.

generally be confined to those of the higher cognitive competences and psychological issues studied or listed predominantly in cognitive science. Publicity and between-individual individuation are not issues that are usually examined or compartmentalized as areas of interest or higher cognitive competences in cognitive science. These topics are important, but they will not be discussed in this essay.

Desideratum #1: Categorization

A prominent criterion for a theory of concepts is the ability to explain categorization. It is largely the need to satisfy this aim for which new concept views have arisen and have been sustained. As will be shown in the next chapter, the prototype, exemplar, and theory-theory views have gained their plausibility in part from this criterion. For natural kinds and three-dimensional artifacts, the criterion of categorization assesses how well concept theories can account for how human beings actually subsume more particular objects and classes within a more general or superordinate category. For example, psychologists may measure how and whether participants classify *bears*, *whales*, and *monkeys* as falling within the *mammal* category. For ethical concepts, one aspect of categorization will be how particular actions are classified as moral or immoral. For instance, a study may measure how and whether subjects classify acts of stealing food when one is poor as falling under *morally right* or *morally wrong*.

As we shall see in much more detail in the next chapter, other aspects of categorization involve typicality effects, where some members of a category are

considered to be more typical than others. More typical members are usually categorized faster than other less typical tokens of a class and are more easily learned. Another matter of categorization is the matching of features between a category and its members. Especially for similarity-based views, members that share more features as well as more heavily weighted features with its category will be deemed by participants to be a more typical token. The ability for a theory of concepts to address and explain these phenomena of categorization add to the usefulness of the theory.

There is also a basic level of categorization where objects tend to be categorized faster and more readily at a particular level of abstraction or inclusiveness rather than at other levels. This is the most natural level where subjects "conceptually carve up the world" by which non-identical category members may be treated as equivalent. Rosch, et al., have discovered that the basic level of concrete concepts tend to be at the middle level of abstraction on a concept hierarchy.[50] For example, in the hierarchy:

Superordinate:	ANIMAL
Basic Level:	BIRD
Subordinate:	ROBIN

Our experimenters claim that BIRD is the middle and basic level of abstraction. Rosch, et al. have discovered that when shown pictures of objects, subject most readily categorize at the middle level as compared to others. Moreover, when asked what a picture of an object is about, participants almost exclusively provide middle level names.

---

[50] E. Rosch, C. Mevis, W. Gray, D. Johnson, and P. Boyes-Braem, "Basic Objects in Natural Categories," *Cognitive Psychology* 8: (1976), 382-439. Eleanor Rosch, "Principles of Categorization.

This basic level is also the first level at which children can sort and name objects. Moreover, children can learn novel basic level categories before those at other levels. As our experimenters claim, the basic level is generally the preferred level of carving up the world. A desideratum of a theory of concepts is to explain this basic level phenomenon of categorization. However, no tests have been run regarding the basic level for moral concepts. Insofar, as it is unclear as to what the basic level is in the realm of ethical concepts, this particular categorization criterion to explain the basic level phenomenon will not be a moral concepts criterion.

Desideratum #2: Cognitive Content

The cognitive content of a concept can be traced back to Frege's view on semantics.[51] Frege described cognitive content, or what he called the sense of a term, as a mode or way of presenting the putative referent. For example, although unicorns do not refer to anything since they are not real, there is still a cognitive content to unicorns that describes the supposed yet non-existent referent. For example, it perhaps may be HAS WINGS, FLIES, and HAS A HORN. Moreover, Frege claimed that an identity statement between two co-referential terms can be informative. For example, the statement "Kierkegaard, who is not the penned author of *The Sickness Unto Death*, is really Anti-Climacus, who is the penned author of *The Sickness Unto Death*," could be surprising

---

[51] G. Frege, "On Sense and Meaning," trans. M. Black, ed P. Geach and M. Black, *Translations form the Philosophical Writings of Gottlob Frege*, (Oxford: Blackwell, 1953).

and educational to many people who lived during Kierkegaard's time. Despite the fact that both names refer to or denote the same man, the referential aspects of the names do not fully exhaust the content of them. Thus, there must be some additional aspect of understanding the terms that is over and above that of reference; namely, the mental state descriptions that one is not and the other is the original formally penned author of *The Sickness Unto Death*. It is this additional understanding that is what we shall call the cognitive content of a term.[52]

Frege also argued that substituting co-referential terms in certain linguistic expressions may alter the truth value of the expression. For instance, "Colleen thinks that Anti-Climacus is the penned author of *The Sickness Unto Death*" differs in truth value from "Colleen thinks that Kierkegaard is the penned author of *The Sickness Unto Death*." For, Colleen thinks or believes that Anti-Climacus is the penned author of the book, but she does not know that Anti-Climacus is the same person as Kierkegaard. She may think that Kierkgaard wrote some philosophy books but at the same time think that he did not write *The Sickness Unto Death*. Thus, both linguistic expressions have different truth values where the first sentence is true, but the latter is false. In spite of the fact that the two proper names refer to the same individual, there is an additional cognitive content to them that accounts for this change in truth value.

---

[52] Frege viewed the sense of a term as abstracta. However, since this view is not held here, we shall use the term "cognitive content." Cognitive content will be understood to be mental states in the head.

Furthermore, Putnam constructs a thought experiment where there is a Twin Earth exactly similar to our own world except in one crucial respect.[53] The people on Twin Earth also have something they call 'water' that is a clear and tasteless liquid and appears to be exactly like the liquid that we on Earth call 'water.' However, the water on Twin Earth is composed of XYZ, not H_2O, like the water on our own planet. Putnam reaches the conclusion that the meaning of the two water terms on Twin Earth and Earth are different since they both refer to different chemical substances. Nevertheless, regardless of whether Putnam is correct that they have different meanings, there is an important respect in which both terms of water are similar. After all, they are both clear and tasteless liquids that fill oceans and rivers. It is these mental state descriptions of both watery substances which are the cognitive content of the terms that allow them to have some similarity in spite of the fact that they may have different meanings and referents.

This example suggests that our concepts also have a cognitive content to them. While concepts are concerned with ways of understanding rather than necessarily drawing any conclusions in semantics about the meaning and reference of linguistic terms, concepts have what Prinz calls a "cognitive content" to them which makes important contributions to how human beings psychologically understand the world.[54] Dating at least back to the Early Modern Philosophers, concepts or ideas are largely responsible for the human understanding. The cognitive content of a concept, when

---

[53] Hilary Putnam, "The Meaning of "Meaning,"" *Philosophical Papers, Vol 2: Mind, Language, and Reality*, by H. Putnam, (Cambridge: Cambridge University Press, 1975).
[54] Prinz, *Furnishing the Mind*, p. 6-8.

discussed at the level of the corresponding linguistic terms, is a way of presenting the purported referent of a term that may allow a statement of the terms' being co-referential to be informative. Also, it may account for the change in truth value of certain linguistic expressions in which two co-referential terms are substituted for each other. Furthermore, cognitive content may allow two terms that are not co-referential to appear to have similarities with each other. Any theory of moral concepts must be able to account for the cognitive content of such concepts in order to explain how human beings psychologically understand the world and use moral concepts in cognition.

As a matter of clarity, the explanation of the cognitive content of a concept has been given in light of the notion of reference. However, providing a theory of reference is not required in order to understand the idea of cognitive content. All that is needed is merely the fact that concepts refer or supposedly refer; we do not here need to discuss how that is so. Regardless of what theory of reference may be correct, so long as concepts have an (putative) extension then we may comprehend the idea of cognitive content in the numerous ways illustrated in this very section. Providing a theory of reference is not necessary in understanding the notion of cognitive content and thus, still remains outside our scope.

Desideratum #3: Acquisition

Theories of concepts must also provide an account of or at least be consistent with the most plausible views on how learned rather than innate concepts are ontogenetically

acquired.[55] For ethical concepts this will primarily be an issue of moral psychology, child development, and how we learn right from wrong. It may be the case that ostension, hypothesis testing, and learning from error, either considered individually or in combination, may be important means by which human beings generally may acquire moral concepts. Acquisition may examine how both children and adults attain their moral concepts by, for example, encountering members of the concept's extension or by being instructed by others.

Desideratum #4: Induction

The induction criterion is where theories of concepts must have a structure that allows and can explain how we make inductive inferences about the world. As induction is a crucial higher cognitive competence that is reliant on concepts, an ability to account for this criterion is important for any theory. There are many different types of induction, but induction in philosophy is commonly understood as a type of reasoning that draws general conclusions from particular cases, where the conclusion does not necessarily have to follow. Philosophical induction was the target of Hume,[56] and it is contrasted with deductive reasoning where the conclusion is necessarily entailed by the premises. An example of philosophical induction in category-based language is drawing the conclusion

---

[55] Here, I do not assume that we have innate moral concepts. Innateness is merely introduced to contrast it with concept acquisition.

[56] To note, Hume's inductive conclusions also included particular predictions that were not general.

that all objects within the category *shark* are dangerous based on previous limited experiences with particular instances of objects that fall under this class. As can be seen, this kind of philosophical induction is fallibly in play in concept acquisition and how we form concepts.

Now, what is meant by "category induction" in the concepts literature is the broadening of information from an already known *given* category to a new action or object that falls within that category, or it may also involve such broadening from a previously known given class to a known but different or *target* category. An instance of the former is that if one has knowledge about *good action*, where acts that fall under this class are ones that a person should perform, and I give you advice in a moral matter that a particular action is good, one may induce that this act is the act that one ought to perform. An example of the latter is where if I am told that those classified under *humble person* have the property of being soft spoken, then I may induce that those under the target class *meek person* also have this property. Category induction will be the focus of this essay as opposed to philosophical induction. How the various viable views of moral concepts handle induction will be discussed in chapter five.

Desideratum #5: Concept Combination

It is thought that concepts have productivity and systematicity.[57] Productivity is where with a finite set of concepts and rules of combination, we can have an infinite amount of new thoughts. For example, if we have the concept AND with the letters of the alphabet, we can have A AND B, A AND B AND C, AA AND B AND BCC, etc. ad infinitum. Systematicity is where if one understands the given concepts and combination rules, one can understand variations on a thought. An example of this is where if one can understand JONES SHOT A FEDERALLY PRESERVED AND ENDANGERED WOLF, then one may also understand the incredible idea A FEDERALLY PRESERVED AND ENDANGERED WOLF SHOT JONES.

Concept combination is the desideratum that attempts to explain how lexical or even compound concepts may combine with each other in order to form even more complex concepts or ideas. Now, compound concepts may still be thought to be concepts. However, for clarification, they will be referred to as "compound" or "complex concepts" in order to mark the difference between them and lexical concepts. Combination supposedly must not only take into account the productivity and systematicity of thought, but it apparently must also account for emergent features, or new components of a compound concept that represent features that apparently "emerge" from a complex class that are not ordinarily thought to be contained within the constituent classes. For example, Smith and Osherson explored the compound concept

---

[57] Jerry Fodor, *The Language of Thought*, (Cambridge, MA: Harvard University Press), 1975.

PET FISH in which this compound concept's constituent concepts, PET and FISH, combine to form the given compound concept.[58] However, PET FISH has as a component LIVES IN A BOWL, but, it is not thought that typical examples of pets, such as a dog or cat, and typical examples of fish, such as bass and catfish, have *lives in a bowl* as a property.[59] Thus, LIVES IN A BOWL represents an emergent feature because it does not appear to be contained in any of the constituent concepts that combine to form PET FISH. A discussion of how the viable theories of moral concepts may account for this criterion will be explored in chapter five.

---

[58] Edward Smith and Daniel Osherson, "Conceptual combination with prototype concepts," *Cognitive Science* 8 (1984), 337-361.

[59] This assumes an exemplar theory of concepts. The exemplar theory will be fully discussed in the next chapter.

2. The Theories of Concepts

This second chapter will introduce the major concept theories. Such an exploration will examine such theories in light of both concrete and ethical concepts, where initial theorizing of how such views will look in the ethical domain will be provided here. The evidence, arguments, and desiderata satisfaction for which theories are viable in the moral domain will be examined in later chapters.

2.1 The Classical Theory

The classical theory of concepts maintains that concepts are composed of bodies of knowledge that are about the individually necessary and jointly sufficient features of members of a class.[1] This view is also known as "definitionism," as it claims that our concepts are bodies of knowledge that represent strict definitions that are used in the higher cognitive competences. Our overview of this theory will be brief since it is generally held to be a discounted view in psychology for concrete concepts, although it is commonly thought to be viable for mathematical and logical concepts. The criticisms of this theory will be provided in the third chapter. Illustrations of the kinds of concepts we have based on the classical theory are that our BACHELOR concept may be constituted by BEING UNMARRIED and BEING A MALE or that an EVEN NUMBER may be made up of A NUMBER DIVISIBLE BY TWO WITHOUT REMAINDER. In the first case, if an individual

---

[1] Clark Hull, "Quantitative aspects of the evolution of concepts," *Psychological Monographs, XXVIII*, 1920. Barbel Inhelder and Jean Piaget, "*The Early Growth of Logic in the Child: Classification and Seriation*, (London: Routledge and Kegan Paul, 1964).

satisfies the necessary and sufficient conditions of being both unmarried and a male, then we will classify this person as a bachelor. In the latter instance, any number that satisfies the definition of being divisible by two without remainder will be categorized as an even number. Therefore, if the classical theory is viable in the moral domain, it may claim one may have in mind a definition such as "an act must maximize the best consequences" when representing the category *right action*.

2.2. The Prototype Theory

In the 1970's, psychologists started running studies which began pointing towards a new theory of concepts. Spearheaded by the likes of Eleanor Rosch, Caroline Mervis, Edward Smith, Lance Rips, and James Hampton, the prototype theory was born as a view of concepts that could account for the new experimental data.[2] Rosch and Mervis explicitly draw inspiration of the prototype view from Wittgenstein's notion of family resemblance, where members of a category may have one to several features or characteristics in common with each other, but zero or very few characteristics are

---

[2] Eleanor Rosch and Caroline Mervis, "Family resemblances: Studies in the Internal Structure of Categories," *Cognitive Psychology* 7: 573-605. Eleanor Rosch, "Natural Categories," *Cognitive Psychology* 4: (1973), 328-350. Eleanor Rosch, "Principles of Categorization," *Cognition and Categorization*. Ed Eleanor Rosch and B. Lloyd. Philadelphia, PA: Lawrence Erlbaum Associates, Inc., 1978. M. Posner and S. Keele, "On the Genesis of Abstract Ideas," *Journal of Experimental Psychology*, 77: (1968), 353-363. E. Smith, E. Shoben, and L. Rips, "Structure and Process in Semantic Memory: A Featural Model for Semantic Decisions," *Psychological Review* 81: (1974), 214-241. James Hampton, "Polymorphous Concepts in Semantic Memory," *Journal of Verbal Learning and Verbal Behavior* 18: (1979), 441-61.

common to all category members.[3] There are various versions of the prototype theory, but it generally states that a concept is composed of a prototype or a body of knowledge that is about a list of statistically frequent features that are summary or general representations of members of a class. Such features are considered statistically frequent in that it is highly probable that a member of one's category will have them. Moreover, prototypes do not represent features that are necessary and sufficient conditions for determining membership. For example, one's prototype of DOG may be HAS HAIR, HAS FOUR LEGS, BARKS, PLAYS FETCH, and WAGS ITS TAIL. The features that are represented may be arrived upon based on all of one's previous experiences with particular dogs. Furthermore, a dog may still be considered a dog even though it is hairless or even if it only has three legs. Such features are not necessary and sufficient conditions. Now, some prototype theories may emphasize that features are based on cue validity rather than statistical frequency, or it may be the case that feature selection of a category is based on both. Cue validity is a measure that increases with the probability that a feature occurs within a category and also decreases with the probability that the feature occurs within members of a similar but different contrast category. For instance, *plays fetch* may be a feature that has high cue validity for *dog* in that while it may be a feature common among the members of *dog*, items in contrast categories such as *wolf* and *hyena* are unlikely to

---

[3] Wittgenstein, *Philosophical Investigations*, trans by G. Anscombe, (New York: Prentice-Hall, Inc., 1953).

have this feature as well. On the other hand, *has hair* does not have high cue validity for *dog* in that it will most likely occur in members of *wolf* and *hyena*.

On some prototype views, based on statistical frequency or cue validity, features may be weighted more heavily than others. For example, *barks* and *plays fetch* may be weighed more heavily than *has hair*. When calculating how similar a potential or target member may be to a category, a prototype theory may take into account the number of features the instance may share with a category, or the instance's satisfaction of heavily weighted features, or both. Prototype theory is considered a similarity-based view because when an object or act is similar enough to the representation of summary or general features and passes a calculated similarity threshold, then the object falls within the class. Returning to the example, since my pet animal satisfies the importantly heavily weighted feature of *barks* and *plays fetch*, along with several of the other features, it passes the similarity threshold and is categorized as a dog. When a token passes the similarity threshold of two or more categories, it is generally categorized in the class towards which it has the highest similarity score. At the same time, passing the similarity threshold for multiple categories can also explain the phenomena of ambiguous cases of membership where some individuals may categorize an item as a member of two different classes.

The prototype view also tends to be a linear model where a category's properties are independent cues for categorization in that the incremental increase of a similarity score occurs when one feature is matched between the potential member and a category

despite the fact that it may not yet have been determined whether there is further feature matching between the target item and category. As we shall see in the non-linear exemplar theory, this cognitive processing contrasts with a non-linear model where the incremental increase of a similarity score is dependent on the sharing of other features between the target item and the category.

Moreover, prototype theorists believe that an item having a higher similarity score to a prototype explains why subjects understand such higher scored items to be more typical members of a category than others.[4] Due to this graded membership or typicality effects, for example, a golden retriever will be thought to be a more typical dog than a Chihuahua for some individuals. Typicality effects usually are measured by asking subjects to rate how typical an item is of a category on a scale. An example of the scale may be: 1=very typical; 4=moderately typical; 7=not very typical. While an item's similarity score does play a role in typicality judgments, where a member with a high similarity score tends to be a more typical member, it has been discovered that other factors, such as a participant's estimate of how many times she has encountered the item, also influence such judgments.[5] Typical members also are usually learned quicker by

---

[4] Eleanor Rosch, "Cognitive representations of semantic categories," *Journal of Experimental Psychology: General* 104: (1975), 192-233. Rosch and Mervis, ibid. Hampton, ibid. S. Armstrong, L. Gleitman, and H. Gleitman, "What Some Concepts Might Not Be," *Cognition*, 13: (1983), 263-308.

[5] Larry Barsalou, "Ideals, central Tendency, and Frequency of Instantiation as determinants of Graded Structure in Categories," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 4: (Oct. 1985), 629-49.

subjects, and tend to be categorized faster as members. Members that are considered to be less typical are more difficult to learn and take a longer time to be categorized.

A common way to examine whether a given concept has prototype structure is by giving participants Hampton's feature listing task or a variation thereof.[6] In the first phase, participants are asked to list what they take to be the features of several categories. In order to tease out the features, several questions are provided to the subjects such as, "why is X only loosely speaking a kind of Y?" or "what characteristics of X make it a Y rather than a Z?" Once the questionnaires are gathered, the features are pooled together to create a much shorter list based on a production frequency measure. Those properties from the pool that score low on this measure are eliminated. The list of features that remain after the elimination round is considered to be the list of features associated with the category. In the second phase, experimenters provide a list of potential category members for each of the previously given categories. Subjects then rate how well they think each potential member belongs or does not belong in the category. Those that are rated highly are taken to be typical members while those that generally are still considered to be members but are not rated as highly are taken to be atypical members. Next judges determine which features of a category from the first phase each member of the category possesses. Given that judges now have determined which features members possess, it may now be seen whether there is a correlation between typical membership and the number and degree of features a member satisfies. It may now also be tested to

---

[6] James Hampton, ibid.

see if the features from the first phase are indeed a prototype or if they are a classical definition. The feature list from the first phase is compared with the category's members' and nonmembers' features derived from the judges. If the concept has prototype structure, then the feature list will not prove to be independently necessary and jointly sufficient conditions for determining category membership and for excluding instances as nonmembers. As we will see in the next chapter, the results of typicality effects, unclear borderline cases of membership, and the absence of psychological evidence for the use of definitions combine to state a case against the classical theory of concepts.

Our examination of the concept DOG presupposes the featural model of prototype theory. This more predominant model for prototype structure uses features for a class in which a potential member is calculated as either having the feature or not having the feature. The featural model does not allow for any middle ground where an object may be thought to only partially have some characteristic to a certain degree. For a featural calculation or cognitive processing, a potential member either has it or does not have it. For example, *has four legs, barks, plays fetch* and *wags its tail* may be thought of as qualities that a potential member of *dog* either has or does not have. Under the featural model, whether it does or does not have such features will in part determine whether or not the token instance may be categorized as a dog.

A second, but less prevalent option as compared to the featural model, as illustrated by Edward Smith and Douglas Medin, is the dimensional model where whether a potential member shares the properties of a category are taken to be a matter of

continuous degrees rather than a matter of bluntly either having the property or not having it.[7] For instance, when using a dimensional model, one may calculate the similarity of particular objects to *dog* based on the dimensions of *size*, *athleticism*, and *ferocity*. The object in question will have different degrees or measurements of these attributes. On this model, it is not an issue as to whether the object does or does not have these qualities. Rather, it is an issue as to what degree the object has such characteristics. At times, it may be the case that certain qualities can appropriately be used both in a featural and dimensional model. On the dimensional model, a prototype will be a body of knowledge that represents the mean or average dimension values for given qualities that have been abstracted from one's experiences of particular instances of a category. Such dimensions, like the featural model, may also be weighted in terms of importance. Thus, whether or not an object passes the similarity threshold of *dog* is a matter of degree based on whether or not the object is close to the average dimension values for the class.

If prototype theory is viable for moral concepts, then the prototype for the moral concept RIGHT ACTION for some individual may be BEING GENEROUS TO OTHERS, HELPING THE HOMELESS IS THE RIGHT THING TO DO BECAUSE IT BENEFITS THOSE IN NEED, PREVENT HARM, DOES NOT BREAK LAWS, and EXHIBITS FRIENDLINESS. Such representations, when understood as features of members of a class, are not necessary and sufficient conditions. As we can see, prototypes may be about such things as general

---

[7] Edward Smith and Douglas Medin, *Categories and Concepts,* (Cambridge, MA: Harvard University Press), 1981.

features of moral situations, virtues, reasons for action, and basic moral principles or rules that happen to occur or be extracted from many particular instances of situations that fall under *right action*. For example, when a person points out to another an instance where a stranger is helping homeless people that such is a case of moral rightness, the mentally represented reason or justification for action that HELPING THE HOMESLESS SO LONG AS ONE IS NOT IN POVERTY ONESELF IS THE RIGHT THING TO DO BECAUSE IT BENEFITS THOSE IN NEED may now be a candidate to be a constituent component of this listener's prototype of RIGHT ACTION based on further particular experiences. The abstraction of such features may be based on personal experiences and moral education. Now, for the individual in question, the virtue friendliness may carry less weight for this individual as compared to heavily weighted features such as the principles *prevent harm* and *does not break laws*. Thus, particular acts such as walking away from a potentially violent physical confrontation with a meddling person or calling the police when witnessing a homicide unfold are more typical instances of right actions than going out of one's way to welcome a new resident to the neighborhood. Now, all of the above given acts supposedly will pass her similarity threshold for *right actions* to be categorized as one of its members. However, as indicated, if moral concepts have prototype structure, there will be graded membership among the acts in question. Furthermore, the more typical members will tend to be categorized faster and learned more quickly as members of *right actions*.

Also on the prototype model, assuming that she lives in a society where abortion is legal, acts such as abortion may be ambiguous for her in that it may barely pass the similarity threshold for both *right action* and *wrong action*. For her, acts of abortion match the heavily weighted feature of *does not break laws,* but it also is a cause of harm in that she believes a fetus has personhood. Given that there are certain hard cases where individuals appear to suffer from moral ambiguity, the prototype theory nicely accounts for and predicts this phenomenon with its given structure and similarity-based calculations.

If moral concepts have prototype structure, a featural model may work for such concepts, where qualities such as *does not break laws* may be considered to be features. However, some characteristics may be more amenable to a dimensional model. For example, the features *exhibits friendliness* and *prevent harm* may be accounted for by matters of degree and a dimensional value rather than by whether a particular act or person does or does not have these features. A person who invites and actively introduces a new, shy, and awkward individual to each person in a social group and attempts to include him in conversation may be considered as friendlier than one who merely says hello on a daily basis to this individual at work. We may certainly think that a good Samaritan has a higher degree of helping as opposed to someone who merely notifies the proper authorities when seeing an injured person on the side of the road. For moral prototype concepts, depending on the list of characteristics a prototype represents, the possibility will be left open that a dimensional model may be the actual way by which

a person may calculate for similarity as to whether a particular act is categorized as moral or immoral. Nevertheless, for our purposes, when discussing the prototype theory we will assume what is considered to be the more predominant featural model, but this is not meant to preclude the possibility of the dimensional prototype model.

2.3 The Exemplar Theory

Soon after the arrival of the prototype theory, some cognitive scientists began hypothesizing the existence of a different theory of concepts. This exemplar theory, like the prototype view, also calculated membership based on a similarity score measure. However, instead of accounting for similarity by comparing the potential member to a summary or general representation of features, it is compared to a set of particular members of a category. An exemplar is a mentally represented particular instance of a class. This view, in which concepts are constituted by exemplars, was initially proposed by Medin and Schaffer's exemplar theory context model.[8] Like prototypes, exemplars are not considered to be about necessary and sufficient conditions for determining membership. There are various versions of the exemplar theory. However, depending on one's exemplar view, an exemplar may itself be composed of a body of knowledge about

---

[8] Douglass Medin and Marguerite Schaffer, "A Context Theory of Classification Learning," *Psychological Review* 85: 1978, 207-38. L. Brooks, "Nonanalytic Concept Formation and Memory for Instances," in *Cognition and Categorization*, ed. by E. Rosch and B.B. Lloyd, (Hillsdale, NJ: Lawrence Erlbaum associates, 1978). D. Medin, and P. Schwanenfluegel, "Linear Separability in Classification Learning," *Journal of Experimental Psychology: Human Learning and Memory* 7: (1981), 355-68.

the set of features or dimensions of an instance, and such features or dimensions may or may not be weighted. Thus, what are stored in long-term memory for DOG are bodies of knowledge that represent particular instances rather than abstracted summary features. For example, one's concept DOG may have exemplar structure. Thus, based on one's experiences, it may be constituted by a set of exemplars such as THAT GREYHOUND ON THE TRACK, MY NEIGHBORS AMERICAN BULLDOGS, MY ROTWEILLER, and MY SISTER'S CHIHUAHUA. One's GREYHOUND exemplar may be composed of a body of knowledge that is about a set of dimensions of the extension of one's exemplar such as FAST, LARGE, and ATHLETIC. One's exemplars for AMERICAN BULLDOG and ROTWEILER is likely constituted by a body of knowledge that represents related values as compared to the dimensional values held by one's GREYHOUND exemplar. On the other hand, one's CHIHUAHUA exemplar will be composed of a body of knowledge that is about quite different values compared to one's other instances of dogs. Thus, in this case, one's concept DOG has exemplar structure and is constituted by a body of knowledge that is about particular instances of a class. Each exemplar has components that are about a set of features or dimensions.

Now, when one comes across an object that is a Doberman pincer, the exemplar theory holds that the target object will be categorized based on its similarity score with one's dog exemplars; a similarity score that must be higher than other similarity scores for contrast categories. Since one's DOG concept is constituted by exemplars and one has exemplars such as MY NEIGHBORS AMERICAN BULLDOGS and MY ROTWEILLER, it is

likely that the Doberman pincer will be categorized as a dog. The set of dimensions for the Doberman pincer will be similar to the sets of dimensions contained by several of one's particular instances. Furthermore, since exemplars do not represent necessary and sufficient conditions, the Doberman will still be classified as a dog even though it comes nowhere close to matching the dimensional values of a Chihuahua. The exemplar theory can also claim to partially account for graded membership, where more typical members are categorized faster and learned quicker than less typical members. For example, given its high degree of similarity to most of one's particular instances, a Doberman pincer may be a typical member of *dog* and may be categorized faster than less typical dogs.

When calculating similarity as specified by Medin and Schaffer's context model, its theory on cognitive processing uses a non-linear model, where the incremental increase of a similarity measure depends on what other features a target item and category share. For example, if there is a pig in my presence, it may have only one instance of feature matching with MY ROTWEILLER concept in that a pig is large. However, a pig is neither fast nor athletic. On a non-linear model of cognitive processing, there is very little to no increase in the similarity score measure, as compared to a linear model used by the prototype theory, since there is only one instance of feature matching. However, if the pig happened to match many more features, then there would be a dramatic increase in the similarity score. A non-linear model allows for the fact that if a target is highly similar to a few or possibly one particular instance, then it is more likely to be categorized as a member of the category as compared to if it is only

moderately similar to all the particular instances that make up the class. This leads to the fact that the exemplar theory is better suited to account for the classification of novel and odd member items as compared to the prototype theory; items that may share few if any properties with a representation of summary or general features. This is a specific respect in which the exemplar theory may be superior to the prototype view. For example, if one comes across a pug, then, with a non-linear model, this will receive a higher similarity score for the exemplar theory since it is highly similar to one's particular instance of a Chihuahua. A pug will be categorized as a dog despite the fact that it does not share many sets of features or dimensions that correspond with one's other instances. However, if one's concept of DOG is a prototype, then at first glance, it appears unlikely that the features or dimensions of a Chihuahua will make one's list of statistically frequent features that are represented by prototypes given the other more predominant experiences one has had of large athletic dogs. Thus, it appears that a pug may not be categorized as a dog on the prototype view because one's prototype may not represent the Chihuahua features or dimensions towards which the pug may feature match.[9] However, due to using a non-linear model, it will be so categorized on the exemplar view.

If moral concepts are exemplars, then an exemplar of, for example, RIGHT ACTION for an individual may be a representation of a grouping of particular cases that one has experienced or heard about that are right acts. For instance, one's exemplars for

---

[9] I say, 'it appears," because the conclusion drawn here may be dependent upon how such things as size and athleticism are weighted.

RIGHT ACTION may be WORKING IN THAT SOUP KITCHEN LAST THANKSGIVING, BREAKING UP THAT FIGHT AT SCHOOL SOPHOMORE YEAR, and MY UNCLE, THE FIREMAN, RUSHING INTO THE TWIN TOWERS ON 9/11. These exemplars themselves have constituent components that are about the features of the extension of the exemplars. For example, WORKING IN THE SOUP KITCEN LAST THANKSGIVING may be composed of HELPING THOSE IN NEED, VOLUNTEERISM, COMMUNITY, and SACRIFICE. BREAKING UP THAT FIGHT AT SCHOOL SOPHOMORE YEAR may be constituted by NON-VIOLENCE, INTERVENTION HELP, COURAGE, and PEACE. MY UNCLE, THE FIREMAN, RUSHING INTO THE TWIN TOWERS ON 9/11 may share similar constituent components with one's exemplar BREAKING UP THAT FIGHT AT SCHOOL SOPHOMORE YEAR. Thus, when one witnesses a policeman intervening in a robbery and assault, since it closely matches the particular instance of breaking up that fight at school sophomore year and one's uncle rushing into the Twin Towers on 9/11, one will categorize this as a right act. Real life examples of the possible use of this kind of exemplars are when precedents are cited in order to determine whether an act is classified as right or wrong.

A second possibility of exemplar theory for moral concepts is that one's concept of GOOD PERSON may be constituted by exemplars that are about individuals who are moral exemplars. In other words, one's moral concept is composed of bodies of knowledge that are about particular individual persons who one takes to be exemplary models of ethical human beings. For example, one's concept GOOD PERSON may be composed of MY MOTHER, MY BROTHER, JESUS, and MOTHER THERESA. These

exemplars that are about moral exemplars in turn each are composed of components that represent a cluster of features. For example, MOTHER THERESA and JESUS may have the components KIND, HELPFUL, LOVING, and BRAVE. When one comes across an instance of a doctor medically helping the poor in a third world country pro bono, since this case is highly similar to one's moral exemplars of Mother Theresa and Jesus, one classifies this person as a good individual. The commonly heard phrase in today's society of "What would Jesus do?" is an excellent example of the possibility of exemplar theory being at work for moral concepts. In this instance, one's concept of GOOD PERSON may be constituted by JESUS, in which individuals are categorized under *good person* if they match the features contained by Jesus. Other such examples are when individuals such as social activists and politicians cite and use particular moral exemplars such as Martin Luther King, Jr. or George Washington when determining which individuals they categorize under *good person*.[10]

---

[10] To note, this possible use of exemplars of exemplars is analogous to the ethical normative theory of neo-virtue ethics where one should perform that action that the virtuous agent will perform. Implicit within this normative theory is that the virtuous agent who exemplifies the virtues; virtues of which lead to eudemonia or a flourishing life, may be an agent from whom it may be determined which acts are categorized as GOOD. The virtues of the virtuous agent in neo-virtue ethics are like the features of the exemplars of the exemplars. Here, particular acts that are similar to the exemplars of the exemplars and match its features will be categorized as GOOD just as in neo-virtue ethics, actions are categorized under GOOD based, in part, on whether they coincide with the virtues that are held by the virtuous agent. Although neo-virtue ethicists are more concerned with the correct concept rather than conception of GOOD, the neo-virtue ethicist's theory to do that action that the virtuous agent will do can be seen to be different yet analogous with this particular aspect of the exemplar theory of moral concepts for the psychological categorization of particular acts under GOOD.

2.4 The Theory-Theory

The theory-theory is a view that emerged out of psychology in the 1980's that became established based on its power to explain categorization as well as on its capacity to provide a detailed account of concept acquisition. Darwin, in *The Origin of Species*, argues that the classification of biological species should not be based on theory-neutral superficial similarities but should be grounded in the causal explanations that underlie the similarities among organisms. Quine in "Natural Kinds" states that an individual's psychological development and the development of society in regards to distinguishing and characterizing natural kinds is first based on perceptual superficial similarities.[11] However, through continual development and maturation, more sophisticated scientific theories are used by the individual and society to draw such distinctions. Along the lines of these thinkers, although not necessarily agreeing with everything they say in their works, the theory-theory of concepts states that concepts are themselves theories or mini-theories. Theories contain scientific, causal, functional, and general or generic background knowledge about the extension of a concept and can explain such things as categorization in concrete concepts.[12]

---

[11] W.V.O. Quine, "Natural Kinds," *Naming, Necessity, and Natural Kinds*, ed. S.P. Schwarz, (Ithaca, NY: Cornell University Press, 1977), 155-75.

[12] Susan Carey, *Conceptual Change in Childhood*, (Cambridge, MA: The MIT Press, 1985). Gregory Murphy and Douglass Medin, "The Role of Theories in Conceptual Coherence," *Psychological Review* 92 (1985). Frank Keil, *Concepts, Kinds, and Cognitive Development*, (Cambridge, MA: The MIT Press, 1989). Alison Gopnik and Andrew Meltzoff, *Words, Thoughts, and Theories*. (Cambridge, MA: The MIT Press, 1997).

A major shortcoming of similarity theories is their inability to account for the additional knowledge that appears to be stored in concepts that accounts for some cases of classification. The theory-theory attempts to address this issue by claiming that background knowledge of the world rather than prototypes that are about superficial properties plays a role in such a process. For example, Francis Keil ran a study where participants were asked whether the animal in a given scenario is a horse or a cow.[13] In the situation, there is an animal that is called a "horse," makes horse sounds, looks like a horse, is strapped with a saddle so people can ride on it, and eats oats and hay. However, scientists run blood tests and x-rays on it, and they discover that its insides are actually the insides of a cow. In this experiment, Keil found that older children and adults perceived the scientists' discoveries as relevant for determining natural kind membership. These subjects relied not on superficial similarities but on folk biological theories of hidden essences to decide that the animal was really a cow despite its superficial horse appearances.

As an example of the importance and use of causal knowledge in cognition, *being curved* is an equally typical feature in bananas and boomerangs. However, subjects give more weight to this attribute in boomerangs rather than bananas because it is falsely believed that curvature is causally related to the boomerang's property of *returning back*

[13] Keil, 1989, 162.

to the thrower.[14] Due to this relationship between the two features, it is thought that *being curved* is more required for a boomerang rather than a banana. Theories may provide other causal explanatory relations between superficial features of an object. As an example that is an oversimplification for the point of illustration, the prototype FISH is constituted by HAS FINS, HAS A TAIL, and SWIMS.[15] Folk theoretical knowledge of fish provides the explanation of the relation between fish attributes since in order to properly swim, a fish needs fins and a tail. Here, theory-theorists typically do not necessarily deny that one may have in mind superficial features when representing a class, but they do emphasize the importance of background knowledge or theories in providing the underlying causal explanation to such features as well as in deciding what weight such features may possess.

Theory-theorists also hold that there are domain differences for types of knowledge where different ontological domains contain different types of central beliefs. For example, while natural kinds are believed to have hidden essences, the analogue for artifact kinds generally is intended function.[16] For example, Lin and Murphy ran an

---

[14] D. L. Medin and E. Shoben, "Context and Structure in Conceptual Combination," *Cognitive Psychology* 20: 1988, pp. 158-190.

[15] Murphy and Medin, 1985.

[16] One may wonder, "Will subjects categorize a broken mousetrap as a mousetrap?" While I am not aware of any experiments testing for this, theory-theory studies, such as the tuk study described below, suggest that if shown a picture of a complete mousetrap, but for some reason the spring in the trap will not work, subjects will still categorize it as a mousetrap since its maker intended the function of it to be a mousetrap. However, if shown a picture of just the flat board of the mousetrap or just the spring of the trap, then participants will not classify it as a mousetrap

experiment where they first described and showed pictures of certain artifacts from foreign countries.[17] One such item is a *tuk*. A tuk is a hunting tool that is a stick with a special handle on one end that protects the wielder's hand from animal bites. On the other end of the stick is a noose that goes around the head of the animal. The function of the tuk is to be able to control an animal by placing the noose over its head. After informing participants about what a tuk is and the function it performs, the experimenters showed participants a picture of what looks like a tuk minus only the special handle. When asked to categorize the item, participants categorized it as a tuk. When shown a picture of what looks like a tuk minus only the noose, subjects did not categorize it as a tuk. This suggests that functional knowledge plays a role in categorization, where participants did not categorize the latter item as a tuk because it could not perform the tuk's function. This contrasts with natural kinds where, for example, so long as an animal still has cow innards or a cow essence, it is still a cow in spite of its superficial appearances of being a horse.

Moreover, theory-theory knowledge may not only contain knowledge of hidden essences, theoretical entities, and causal knowledge, but they may also contain knowledge of explanation and prediction. For example, Murphy and Medin claim that

---

because such single items when presented alone do not have the intended function of being a mousetrap. Carey, 1985. Keil, 1989.

[17] E. Lin and G. Murphy, "The effects of background knowledge on object categorization and part detection," *Journal of Experimental Psychology* 50A: (1997), 25-48.

most people think the feature of *flammable* is a quality of wood rather than paper money even though both wood and paper money are flammable. The reason behind this is that we have general knowledge about the world concerning human activity where wood is used for burning fires and paper money is mostly used for economic purposes in which its flammability plays no role. On their view, this knowledge still counts as a theory that influences feature attribution for classes. For, such knowledge in principle does provide an explanation of why we may attribute certain features to certain classes.

A philosophical matter that deserves attention is the actual relationship between theories and concepts. On the one hand, concepts are at times understood to be theories or mini-theories. It is also generally held that theories causally influence concepts. Here, concepts are not themselves theories, but concepts are affected by theories. However, Prinz correctly notes that if this is the case, then this provides no explanation of the structure of concepts.[18] Thus, Prinz proposes that for the theory-theory, the components of a concept are knowledge of hidden essences. Such knowledge of hidden essences may represent necessary conditions that may be empirically defeasible given further findings in science, so therefore, such knowledge may be understood to occupy empirically defeasible placeholder positions. The fact that hidden essences may represent necessary conditions is empirically supported by Barbara Malt's famous study on WATER.[19] When asking subjects to list what percentage of various kinds of water, like purified water,

---

[18] Prinz, 81-82.

[19] Barbara Malt, "Water is not H_2O," *Cognitive Psychology* 27: (1994), 41-70.

swimming pool water, and swamp water, are H_2O, positive but varying percentages were given to all forms of water presented, as well as to several items that are not typically considered to be water. Thus, Malt concludes, "It should be noted, though, that H_2O was judged to be present in all the water examples collected. It may be that one property, the presence of H_2O, is necessary, but not sufficient, for a liquid to be considered water."[20] Unlike definitionism, as a psychological matter, such knowledge of hidden essences does not represent explicit necessary and sufficient conditions. Furthermore, constituent components may be knowledge of explanatory relations that are acquired from background information rather than superficial appearances. In this case many concepts are mini-theories. Second, concepts can be constituents of mini-theories, where some constituent concepts are also bottom level primitive concepts. These primitive concepts are not mini-theories and have no structure. If concepts have other concepts as constituents, then these constituent concepts also have constituent concepts, ad infinitum. In order to stop the regress, theories of concepts need to posit primitive concepts. Third, mini-theories may be causally influenced by other mini-theories. For instance, one's mini-theory of AVIATION can cause one to attribute the knowledge WINGS ENABLE FLIGHT to one's mini-theory of BIRD.

Adopting and applying Prinz's clarification of the theory-theory to ethical concepts, ethical concepts may themselves be mini-theories that have as components such things as knowledge about master moral principles from which other basic moral

---

[20] Malt, ibid., 66.

principles are derived. For example, the concept RIGHT ACTION may be composed of normative theoretical information akin to divine command theory. The theory ACT IN ACCORDANCE WITH THE PRINCIPLES MANDATED BY GOD is a component of RIGHT ACTION. The full explanation of "master moral principle" will be provided in the subsequent paragraph. Theory-theory components may be acquired from background information and may be considered to occupy defeasible placeholder positions rather than strict necessary and sufficient conditions. I do not claim that moral theory knowledge represents necessary conditions, as with hidden essences, because PAC requires that empirical evidence be presented in order to substantiate such a claim, although I do leave the door open that future research may establish this.[21] Second, some concepts will be constituents of mini-theories where some constituent concepts inevitably will be primitive concepts that are not mini-theories and that will allow ethical concepts to escape the infinite regress. Third, it is possible that some ethical mini-theories may causally influence other mini-theories. For instance, one's mini-theory of JUST may influence what features one thinks are attributed to *political policy*.

As previously stated, prototypes are about features that may be moral principles, reasons for action, and virtues. One may derive such principles, reasons, and virtues as

---

[21] In the fourth chapter, we will explore an example in which theory knowledge does not represent necessary conditions. One may hold two different theory beliefs with the same equally strong conviction, and these two beliefs may produce contradictory judgments on a particular matter. In this case, the subject holds two contradictory judgments, where neither of the two theory knowledge is taken to represent a necessary condition that applies to both contradictory judgments.

general or summary representations from particular moral situations or individuals. However, the theory-theory components are more about master or top moral principles from which those middle moral principles, reasons, and virtues may be thought to purportedly derive. Drawing on a distinction made by Robert Audi, master or top moral principles are principles from which middle principles may be inferred.[22] For example, from divine command theory, where one must obey those laws mandated by God, one may arrive upon middle principles such as *do not lie* and *do not steal*. One may adhere to these middle moral principles based on one's background belief in the ethical theory of divine command theory. While the theory-theory is a distinct theory from the others based on the specific kind of knowledge it posits concepts as containing, as we can see, there is an intimate link between the theory-theory and the prototype theory for moral concepts. This is just like the intimate link between both concept theories in the concrete realm, where as shown in our previous example, background explanatory knowledge is the reason why the folk attribute the prototypical property of flammable to wood but not paper money.

It is knowledge about these top principles that constitute the theory-theory of moral concepts, while the middle moral principles, reasons, and virtues belong more in the domain of the prototype view. The reason why this is the case is that such top principles are those that ostensibly underlie the middle principles, virtues, and reasons

---

[22] Robert Audi, *The Good in the Right,* (Princeton, NJ: Princeton University Press), 2004.

for action at a deeper level of theoretical abstraction just as, for example, biological theories underlie the superficial features of a biological natural kind. Second, and a related point, is that just as biological theories may be about explanatory relations between superficial features, top moral principle knowledge may provide an explanatory link between middle principles, virtues, or reasons. In this respect, middle principles, reasons, and virtues may be thought of as being superficial, although there may be varying levels of superficiality at this middle level. For instance, the middle principle *do not shoot innocent persons* can be explained by the less superficial middle principle *do not kill*. Meanwhile, top principles are more theoretical and lie at the deepest explanatory level. For example, *do not kill* may be ultimately explained by divine command theory. Now, based on potential further theorizing, what I have provided thus far for an explanation of the structure of the theory-theory for moral concepts may not exactly parallel the structural components of the theory-theory for concrete concepts given the differences between the abstract and concrete domains. However, in allying some theories with knowledge of top moral principles, we can see that several important similarities as just discussed exist between the given moral and concrete theory-theory structures to warrant drawing the distinction between moral concept prototype and theory-theory structural components in this manner.[23]

---

[23] There is a theory of concepts called the ideals view that is commonly thought to be a specified version of the theory-theory. It states that concepts are constituted by mental representations that are about features that members of a class ought to have to determine membership. Moreover, such components are goal oriented background knowledge. A class of concepts that is thought to

2.5 The Emotion Theory

While there are numerous theories of concepts in the concepts literature, very little attention within this field has been paid to the view that concepts may be composed of sentiments and emotions. However, I believe that some of our concepts may be so constituted and that scientists and philosophers of the mind need to have greater awareness of this fact. This theory at hand will be called the *emotion theory*. It claims that some moral concepts are sentiments and emotions.[24] Before discussing the nuances

---

have ideals structure is ad hoc concepts. Ad hoc concepts refer to novel classes that are not usually stored in long-term memory. They are also instrumental in achieving goals. Some examples of what are thought to be ad hoc categories are *things to take from one's home during a fire*. For example, when determining what to take from one's house when it is burning, some things might be pictures, computer, contact lenses, money, children, and pets. The smorgasbord of members may not lend ad hoc classes to have a set of superficial summary features that are closely correlated with each other. Rather, there is most likely some kind of abstract goal-based knowledge that is at work when determining membership in such classes. It is such background features that the ad hoc category ought to have that may account for typicality effects. This contrasts with potential other concrete theory-theory concepts in which the components lack a normative flavor and are not goal-oriented knowledge. I do not further discuss this view since it is merely a specific kind of theory-theory and furthermore, it is questionable whether the components of moral concepts represent features that members of a class should have. For example, in my GOOD LIFE concept, components of it, such as ATTAINMENT OF HAPPINESS, seem to represent features that simply belong to some lives that are good. It seems odd to say that such features are ones that good lives ought to have. However, even if the relevant knowledge components of moral concepts represent normative and goal oriented features, then the ideals theory for moral concepts is still just a kind of theory-theory for moral concepts, and we need not expound further into this view.

[24] This view differs from the others previously discussed since here, concepts are actually constituted by emotions. While other views such as the prototype theory may mention emotions by claiming, for example, that RIGHT ACTION is composed of BRINGS HAPPINESS, ELICITS APPROBATION, and ELICITS JOY, the emotion theory may be thought to *use* rather than *mention* emotions since moral concepts are actually constituted by them rather than being

of the emotion theory, it will be best to first discuss Prinz's neo-empiricist theory of moral concepts since the emotion theory is very similar to it. After detailing Prinz's view, the emotion theory will be distinguished from Prinz's neo-empiricist theory in that the emotion theory does not necessarily espouse neo-empiricism.

For his neo-empiricist view of moral concepts, Prinz states, "Thus, we can capture the idea that moral concepts are perceptually-based detectors of moral properties by postulating that moral concepts are constituted by sentiments."[25] In *The Emotional Construction of Morals* Prinz attempts to show that moral concepts are constituted by sentiments and emotions, and his method for doing so will later be examined and scrutinized in chapter four. [26]

---

composed of concepts that refer to them. Emotion theory structured concepts sharply differ from the others in that they contain different information carrying mental states (emotions) as well as in the fact that they rely on emotion-based processing. In that emotion theory concepts have a cluster of properties or scientific generalizations that differ from other types of bodies of knowledge, such concepts may be thought to constitute its own natural kind.

[25] Ibid, 97.

[26] Prinz's essentially Humean view of moral concepts is also somewhat mirrored by meta-ethical non-cognitivists who also generally espouse a Humean theory of moral concepts. For, meta-ethical non-cognitivism generally claims that moral judgments are the expression of one's emotions. Insofar as moral judgments are the expression of one's attitudes, meta-ethical non-cognitivists generally are committed to the view that the moral concepts contained within the judgment are constituted by emotions. It is fair to say that the concepts literature in the Philosophy of Mind/Cognitive Science has largely ignored or has been unaware of meta-ethical non-cognitivism's stake on an emotion-based theory of concepts. On the other hand, I also believe it is fair to say that meta-ethical non-cognitivists have largely ignored or have been unaware of the many nuances and theories that have developed in the concepts literature; issues such as the compositionality of thought. Due to the more sophisticated nature of Prinz's theory

First, neo-empiricism as a theory of concepts will be explained with most of the emphasis being placed on Prinz's version of this view. It will then be illustrated how there is an apparent contradiction in Prinz's theory of moral concepts between his first book *Furnishing the Mind* versus his later writings in *The Emotional Construction of Morals*. His initial effort can be seen as a structurally dualistic theory of moral concepts, where such concepts may have exemplar and sentiments/emotions structure, but his later work appears to adopt a monistic view where moral concepts are only constituted by sentiments/emotions.[27] His earlier work is a dualism in the sense that there are the dual exemplar and sentiment/emotions structures for moral concepts. Next, both views will be reconciled such that Prinz's theory may still be construed as a dualism. Now, remember that the emotion theory claims that some moral concepts are composed of our sentiments and emotions. Hence, the emotion theory may appear to be similar to Prinz's sentiments/emotions structure for moral concepts. However, in the final part of this section, these two views will be differentiated from each other.

Neo-empiricism finds its roots from the British empiricist philosophers in the 17th and 18th centuries such as Locke, Berkeley, and Hume. However, while the British empiricists, in varying degrees, generally viewed concepts as conscious picture-like images that are derived from perceptual states or what Hume calls "impressions," neo-

---

of moral concepts at the concepts level, we will focus on Prinz's Humean account of moral concepts rather than meta-ethical non-cognitivists' Humean account. Prinz, *Emotional*.

[27] Prinz's distinction between sentiments and emotions will shortly be given.

empiricists like Barsalou, on the other hand, believe concepts are records of information carrying neural states that underlie perception.[28] Barsalou believes that such neural representations are unconscious though they may make conscious counterparts. While there are different versions of neo-empiricism, they generally hold the view that information in concepts are encoded in modal or perceptual symbols that are represented in the same systems as the perceptual states that produced them. For neo-empiricists, perception is conceived of broadly to include audition, vision, touch, smell, taste, and emotions. The view that emotions are perceptions will be explained shortly. Furthermore, such theorists believe that the higher competences involve the use and manipulation of such perceptual symbols, where conceptual processing concerns the reenacting and manipulation of some perceptual states. Prinz states, "Tokening a [concept] is generally tantamount to entering a perceptual state of the kind one would be in if one were to experience the thing it represents."[29]

This is contrasted with the more predominant view that concept knowledge is stored in arbitrary amodal or non-perceptual formats. For amodal symbol systems, once a perceptual state arises in the sensory-motor system, it gets transduced or converted into

---

[28] Lawrence Barsalou, "Perceptual Symbol Systems," *Behavioral and Brain Sciences*, 22, (1999), 577-660. L. Barsalou, W.K. Simmons, A. Barbey, and C.D. Wilson, "Grounding conceptual knowledge in modality-specific systems," *Trends in Cognitive* Science 7 (2003), 84-91. L. Barsalou, D. Pecher, R. Zeelenber, W.K. Simmons, S.B. Hamann, "Multimodal simulation in conceptual processing. *Categorization inside and outside the lab: Essays in honor of Douglas L. Medin*, ed. W. Ahn, R. Goldstone, B. Love, A. Markman, Nad P. Wolff, (Washington, D.C.: American Psychological Association), 2005.

[29] Prinz, *Furnishing*, 150.

an amodal symbol system to be used in the higher cognitive functions. Amodal symbols are language-like in that just as the word 'dog' is not similar to a physical dog, the amodal DOG symbol is not similar to a perceived dog. Moreover, just as processing a language requires the processing of individual words in a sentence, conceptual processing is based on the processing of individual amodal symbols that are in sentence-like or list-like structures.[30]

While Locke motivated his empiricism from an anti-nativist stance, neo-empiricists need not be committed to anti-nativism. Rather, they may motivate their view from alternative means. Two major forms of motivation for neo-empiricism and the positing of modal symbols are the issues of parsimony as well as numerous findings in neuroscience and cognitive science that indicate that concepts are stored in modal rather than amodal formats. For example, Barsalou and Prinz support their neo-empiricism based on an argument from parsimony. Remember that amodal theories posit that modal representations get transduced into amodal representational formats for cognition. However, it will be more cost-effective if our modal representations can be used in cognition without the use of amodal symbols. If we already have modal symbols and they can do the proper work in all the various facets of cognition, then it will be more parsimonious not to posit the existence of amodal symbols. Moreover, neo-empiricism

---

[30] Jerry Fodor, *The Language of Thought*, (Cambridge, MA: Harvard University Press, 1980). Jerry Fodor and Zenon Pylyshyn, "Connectionism and Cognitive Architecture: A critical analysis," *Cognition* 28: (1988), 3-71.

may be independently supported based on empirical data.[31] For example, McCarthy and Warrington examined patients with focal brain lesions and found that such subjects lose their ability to categorize in certain domains.[32] For instance, in some cases patients' categorization abilities are impaired for abstract concepts rather than concrete concepts. The explanation offered is that concepts consist of modality-specific perceptual information and, in that different category domains may be represented by different sensory representations, such selective impairments will occur. To note, while this modal versus amodal issue has been introduced in order to provide a sufficient explanation of the neo-empiricist view of concepts, no stand will be taken on this debate as to whether all moral concepts are stored in modal or amodal representational formats, as this issue concerning the vehicles of thought is outside the scope of this essay and may itself comprise an entire dissertation.

Since Prinz discusses moral concepts with more depth than any neo-empiricist, his brand of neo-empiricism will be the central view for our discussion. In that it takes Hume the three books that comprise his *Treatise of Human Nature* – "of the Understanding," "of the Passions," and "of Morals" – to expound his view of moral concepts, it may be read that it also takes Prinz his first three books on the same

---

[31] Antonio Damasio, "Time-Locked Multiregional Retroactivation: A Systems-Level Proposal for the Neural Substrates of Recall and Recognition," *Cognition* 33: 1989, 25-62. Barsalou, Perceptual Symbol Systems. D. Morrow, S. Greenspan, and G. Bower, "Accessibility and Situation Models in Narrative Comprehension," *Journal of Memory and Language* 26: 1987, 165-87.
[32] R. McCarthy and E. Warrington, *Cognitive Neuropsychology: a Clinical Introduction*, (San Diego, CA: Academic Press), 1990.

sequentially presented general subject matters of concepts, emotions, and ethics to make his broadly Humean claim on the nature of moral concepts. We will first examine his view in light of his first book, *Furnishing the Mind*. In this work he endorses *concept empiricism*:

> **Concept Empiricism:** all (human) concepts are copies or combinations of copies of perceptual representations.[33]

Here, copying is to be understood as a causal process, where perceptually produced representations may be duplicated in other systems or where perceptual representations leave behind records, which are instructions for producing copies, in other non-perceptual systems that allows for regeneration in the original perceptual systems. Prinz believes that perceptual representations are in dedicated input systems that use different kinds of mental representations for different sense modalities. An input system is a system that receives input from outside the brain in the external environment or from within the body such as through proprioception or interoception. An input system is dedicated in that the different senses are tuned to disparate physical magnitudes. For example, vision responds to wavelengths of light while olfaction is sensitive to molecular shapes. Furthermore, for Prinz, an input system is dedicated in that different modalities have different representational primitives. This is based on such reasons as 1) that our senses have different types of information processing, 2) different types of representations may be better able to handle different sensory-based tasks, and 3)

---

[33] Prinz, *Furnishing*, 108.

our sensory modalities are relatively independent systems. The concept empiricism thesis and the fact that perceptual symbols are in dedicated input systems leads Prinz to the Modal-Specificity Hypothesis:

> **The Modal-Specificity Hypothesis**: Concepts are couched in representational codes that are specific to our perceptual systems.

Prinz calls perceptual representations "proxytypes" since such representations stand in as proxies for the categories they represent. "Proxytype theory" will be understood as Prinz's specialized neo-empiricist view of concepts. Thus, proxytypes are perceptually derived representations that are or can be activated in working memory to represent a class. Long-term memory networks, or a group of linked proxytypes, can store a wide variety of information that concepts are thought to store. Proxytypes may accomplish this feat because they may be multimodal, a single visual model, or a representation of a word, such as an auditory representation of "dog." Prinz claims that proxytypes are bodies of knowledge that represent summary features, particular instances, and theoretical causal explanatory facts that are acquired empirically. For example, one's concept of DOG may be constituted by a prototype that contains BARKS, PLAYS FETCH, HAS FOUR LEGS, and HAS A TAIL. Proxytypes may also be exemplars that are about mentally represented particular instances such as the exemplars THAT GREYHOUND ON THE TRACK, MY NEIGHBOR'S AMERICAN BULLDOGS, MY ROTWEILLER, and MY SISTER'S CHIHUAHUA for one's DOG concept. Proxytypes may also contain causal explanatory knowledge of the world. Thus, Prinz's view is a pluralism which includes the prototype, exemplar, and theory-theory, although proxytypes are clearly

taken to be modal representations.[34] Finally, Prinz states that his view on moral concepts may be understood as a conception rather than correct concept view.[35] Given that Prinz's primary focus is on moral "concepts" as being mental representations in the head and that he is not talking about "concepts" as being abstract objects that contain the universally correct means for categorizing acts as moral or not, Prinz may be read as holding a conception rather than a correct concept theory.[36]

In *Furnishing the Mind* and specifically in relation to moral concepts, Prinz argues that ethical concepts such as VIRTUE may be constituted either by emotions or by exemplars that are about specific acts of moral deeds and good actions. Prinz believes moral concepts may have exemplar structure based on his personal observation, rather than on evidence consistent with PAC, that when asked about the nature of our moral concepts, we will eventually provide examples of particular moral cases. The claim that

---

[34] Prinz states that "Proxytype theory is a *hybrid*…" But it is more charitable to read him as a pluralist since he does not maintain that there is a criterion of correctness that may be competently mastered. Prinz, *Furnishing*, 164. My italics.

[35] Based on conversation with Prinz and Prinz, *Emotional*, p. 94.

[36] Prinz also holds a moral subjectivism, where moral truth is relative to the individual and where there is no universally correct way for categorizing acts as moral or not. To note, Prinz still believes that moral concepts and judgments refer and have truth conditions. For, on his view, moral concepts that are constituted by sentiments and emotions may represent both secondary qualities and concerns via an informational and functional theory of content for emotions. Having truth conditions is perfectly consistent with focusing one's discussion on moral conceptions or moral mental representations. The main emphasis as to why Prinz's view is a conception rather than correct concept theory is that he understands "concepts" to be in the head and constituted by emotions rather than being abstracta outside the head. Second, when discussing "concepts," he is not talking about a morality where there is a universal moral truth since he is a subjectivist. For Prinz, moral conceptions can still refer to secondary qualities and concerns and have truth conditions even though he believes that there is no such thing as an abstract universally correct moral concept.

concepts are constituted by emotions is held due to the work done by Antonio Damasio and how mental representations of actions are accompanied by "somatic markers." These markers cause us to experience certain emotions when a particular act is brought up in cognitive deliberation as to whether we should perform the act or not, where the emotions are thought to play a role late in moral deliberation. When a mental representation about a particular act is accompanied by a positive marker, we will be motivated to perform the action. When negative, the relevant somatic marker leads us to avoid the action. Given that somatic markers may ground mental operations in moral cognition and practical reasoning, moral concepts may be constituted by emotions. While his focus for ethical concepts is primarily on these two views in *Furnishing the Mind*, when taking into account the other previously discussed structures such as the prototype and theory-theory that proxytype theory is compatible with, it may well be possible that Prinz may also include such other structures as being viable for moral concepts as well. Although this possibility will be left open, our focus on Prinz's moral concepts view in *Furnishing the Mind* will be on his explicitly discussed exemplar and emotion-based structures for moral concepts. Before moving on, more will now be stated in regard to the emotion structural aspect of his view.

In his second book, *Gut Reactions*, Prinz maintains the James-Lange view that all emotions are perceptions of our bodily states rather than action tendencies that dispose us

to act.[37] When a bully approaches a student, the oncoming bully causes the student to perspire, tremble, and have an elevated heart rate. On this view, the emotion of fear is merely the experience or perception of these bodily states rather than that which disposes us to perspire and tremble. It is from such perceptions that the qualia that constitute emotions are produced. In other words, we feel fear because we tremble rather than tremble because we are afraid. Insofar as emotions are felt perceptions, neo-empiricism may account for moral concepts being constituted by the passions since proxytypes are perceptual representations. Thus, our first interpretation of Prinz's theory of moral concepts is that of a structurally dualistic neo-empiricist view where our moral concepts are composed of exemplars and emotions which are stored in perceptual representation formats.

There is a second possible interpretation of Prinz's theory of moral concepts that apparently contradicts the first interpretation. In his later book, *The Emotional Construction of Morals*, Prinz holds what he calls *epistemic emotionism*, which contends that our moral concepts are essentially related to emotions. By "essentially related," Prinz means that moral concepts may be constituted by dispositions to feel emotions and, at times, be made up of emotions themselves. Prinz distinguishes between sentiments and emotions, where sentiments are standing dispositions stored in long term memory to feel emotions. Emotions are occurrent manifestations in working memory of

---

[37] Jesse Prinz, *Gut Reactions*, Oxford: Oxford University Press, 2004.

dispositional sentiments. In making this distinction, Prinz allows for the fact that at times one may make a moral judgment without feeling any emotions, but one is still disposed to feel such emotions. Thus, in such cases of judgment, we may say that our moral concepts, which are, in part, the building blocks of moral judgments, are themselves constituted by our sentiments, which are dispositions to feel emotions of approbation and disapprobation even though one does not presently experience any emotions. In other cases where one does feel emotions when making a moral judgment, we may say that one's moral concept in the moral judgment when activated occurrently in working memory is constituted by emotions. All moral concepts which are stored in long term memory are composed of sentiments, but when a moral judgment is actively rendered in working memory, the moral concept in the judgment in many but not all cases is constituted by emotions.

Prinz appears to waver at times in *The Emotional Construction of Morals* as to whether moral concepts are constituted by sentiments or if they are, on the other hand, constituted by emotions. However, as we shall see shortly below in Prinz's story of how we come to make moral judgments, in certain passages it appears that he says that moral concepts are constituted by sentiments and at times by emotions. I adopt this latter interpretation of Prinz since it is able to account for the nature of moral concepts in both long-term and working memory. As concepts are stored in long-term memory and recruited into working memory when actively forming thoughts and judgments, moral concepts on Prinz's view can have different primary natures depending on whether the

moral concept is in long-term or working memory. This variance allows for my previous interpretive statement that moral concepts may at times be constituted by emotions. When simply stored in long-term memory and not being actively rendered, a moral concept is a disposition. However, when the concept is brought into working memory such that it is a constituent of a moral judgment and the disposition is manifested into an actual experience of an emotion, then it is primarily the case that the occurrent moral concept is constituted by emotions since the judgment of which the concept is a constituent is an expression of an emotion. To note, Prinz may maintain that there is still some kind of disposition to feel emotions in this occurrent state – for example, one may still be disposed to feel the emotion more intensely or for a longer period of time. Therefore, in order to account for the different possible stages a moral concept may be in, I take it that a comprehensive interpretation of Prinz's view will be that moral concepts are constituted by sentiments and at times by emotions. Hence, we may say that Prinz's epistemic emotionism claims that moral concepts are related to emotions dispositionally and at times constitutionally.[38]

In holding epistemic emotionism, Prinz does not explicitly state that our moral concepts may be constituted by other things such as exemplar features. He only states that our moral concepts are constituted by sentiments and at times by emotions. Insofar

---

[38] Notice also that this interpretation is consistent with his statement in *The Emotional Construction of Morals* that since sentiments represent secondary qualities and emotions represent concerns, when an emotion is experienced during judgment, the moral concept represents both a secondary quality and a concern.

as this is the case, it initially appears that no other concept structure may be allowed in Prinz's more mature view to allow for the type of moral concept dualism maintained in his earlier work. In *Gut Reactions*, he holds a purely non-cognitivist rather than cognitivist theory of emotions, where emotions have a felt and perceptual aspect but do not contain cognitive appraisal judgments and are not composed of the relevant concepts; concepts which may be the constituents of appraisal judgments.[39] Here, a cognitive state is a specific kind of propositional mental state made up of a collection of individual semantically interpreted concepts. A cognitive state is more precisely an appraisal judgment. An appraisal judgment is an evaluative judgment related to one's well-being. For example, if someone strikes me in the face, then I experience the emotion of anger. If my anger is cognitive, then such anger may be made up of a collection of highly structured appraisal judgments related to such things as goals, blame, praise, coping potential, and future expectations. Such anger may be made up of MY WELL-BEING HAS BEEN DAMAGED, HIS ATTACK IS UNPROVOKED, and HIM HITTING ME IS PAINFUL.[40] However, for Prinz's pure non-cognitivism, moral emotions do not contain cognitive appraisal judgments of things like moral blame or praise such that these cognitive appraisal may themselves be some other concepts structure like exemplars that are always

---

[39] A deeper discussion of pure non-cognitivism will be discussed below.

[40] As judgments are conclusive thoughts that something is in fact the case and concepts are the constituents of thought, judgments are made up of a string of connected concepts. In this cognitivist situation being discussed here, a concept may be constituted by emotions, where emotions are just appraisal judgments. Therefore, the concept in question is constituted by appraisal judgments or a string of relevant concepts. This concept structural formation is similar to, for example, the prototype or theory-theory views.

attached to the emotions. Emotions are not tied to cognitive states. Thus, Prinz's later writings on moral concepts appear to be a monistic view in that moral proxytypes are constituted by our sentiments/emotions *tout court*. It may not be the case that our moral proxytypes are also constituted by exemplars. At first glance it seems that if our moral concepts are constituted by our sentiments/emotions, then moral concepts are simply constituted by our disposition to have pure feelings and perceive bodily changes as well as at times to be constituted by such feelings and perceptions themselves with no other conceptual structures at work. In other words, his more mature and focused view appears to be a monistic rather than dualistic theory of moral proxytypes.[41] Moreover, he makes no further claim in *The Emotional Construction of Morals* that other concept structures may also be at play.

However, with some additional labor, Prinz's disparate dualistic and monistic views in his earlier and later works, respectively, may be reconciled. In order to accomplish this task, we will examine Prinz's story of how one comes to make moral judgments. This story will have a certain opening where it may be plausible to add in the

---

[41] Prinz's later view in *The Emotional Construction of Morals* may be construed as a structural dualism in that moral concepts are constituted by sentiments and at other times emotions. Sentiments and emotions may be viewed as two different things. However, for our purposes, what is important is that his later view does not initially appear to allow for other predominant concept structures such as the exemplar structure that are discussed in his earlier work. It is with this apparent discrepancy in mind that I state that his later work is a monism rather than a dualism. While it is possible that his later view may still at first glance appear to be a dualism, my use of the term "monism" and "dualism" is meant to indicate the apparent discrepancy between his works, and these terms may be altered in a qualified way without loss of effect if need be.

exemplar structure to moral concepts. In this story, Prinz asks us to imagine that a person witnesses an act of theft. Prinz claims that first this person categorizes the act as an act of theft. After this, in long-term memory, one has a sentiment towards theft that becomes activated. In that a sentiment is kind of like a standing belief, Prinz calls sentiments "moral rules." The activated sentiment along with contextual factors then determines what emotions become elicited. Let us say that in this particular case, one has elicited the emotion of anger at theft. The state of being angry at theft constitutes the judgment that theft is wrong insofar as the emotion is generated by a sentiment. Prinz states that in that an emotion is somewhat like an "occurrent belief," only in the limited sense that the emotion is what constitutes the occurrent judgment, Prinz calls relevant moral emotions "moral judgments."

Now, since sentiments are dispositions that generate emotions rather than being non-cognitive emotions themselves, we may claim that sentiments are, in part, constituted by exemplars such as particular instances of theft, robbery, and pick pocketing. We may now claim that in Prinz's story, once an act is determined to be one of theft, one's sentiment towards this act becomes activated precisely because one's sentiment or moral rule is partly constituted by exemplars of stealing. This introduction of exemplar theory adds to Prinz's story in that it accounts for why a particular sentiment is activated when a certain act is witnessed or discussed. The activation occurs because the particular act in question matches the exemplars contained in the sentiment. The activation of the sentiment then leads to the elicitation of emotions and eventually moral judgment. In this

fashion, we may reconcile the disparate views of concepts contained in Prinz's earlier and later works. Since epistemic emotionism claims that 1) moral concepts are constituted by sentiments and at times emotions; 2) sentiments are in part composed of exemplars; and 3) emotions are pure perceptual-based feelings, it may be thought that overall, Prinz consistently holds a structurally dualistic view of moral proxytypes across his disparate works.

At this point, we will now discuss the difference between the emotion-based aspect of Prinz's reconciled proxytype dualism versus the emotion theory. The James-Lange view of emotions as perceptions is highly contentious, where it does appear that the James/Lange theory is the minority view. An example of the controversial nature of this debate can be witnessed by Joshua Greene's work in moral psychology in which he presupposes the view that emotions are not perceptions.[42] Nevertheless, this philosophical debate on the nature of emotions is largely beyond the scope of the dissertation. Therefore, as mentioned, a general emotion based view for moral concepts will be adopted and will be called the emotion theory. The emotion theory differs from Prinz's emotion-based aspect of his neo-empiricism in that it is able to remain theory neutral concerning what emotions are, and as we shall see in later chapters, it is

---

[42] To note, in the *Emotional Construction of Morals*, Prinz claims it may be the case that emotions may be both perceptions and action tendencies. If this is the case, Prinz can still hold onto the view that concepts are modal symbols since emotions are still in part perceptions. Joshua Greene, "The Secret Joke of Kant's Soul," *Moral Psychology Volume 3*. Ed Walter Sinnott-Armstrong. (Cambridge, MA: The MIT Press), 2008.

established and defended by new means.[43] Emotion theory moral concepts are not necessarily perceptual-based representations because I leave open the possibility that emotions are not perceptions. The emotion theory is not necessarily a neo-empiricist view. Hence, we shall entitle the relevant moral concepts view the emotion theory, in which moral concepts are constituted by sentiments and emotions. To note, it will continually be maintained that some moral concepts may be constituted by sentiments, which will still be understood to be dispositions to feel emotions. On this view, moral judgments need not necessarily involve occurrent emotions if the moral concept in the moral judgment is only constituted by sentiments.

At this juncture, we will examine several objections against the emotion theory. Objections against other concept theories such as the prototype view will be anticipated in later chapters. This is due to the fact that such objections pertain to specific desiderata of theories of moral concepts, such as concept combination. As discussed in the first chapter, discussion of such desiderata has been assigned to later chapters. Therefore, such objections will be entertained later. However, I anticipate in this chapter several objections against the emotion theory that are not pertinent to the stated desiderata of theories of moral concepts as outlined in the first chapter. The first objection is concerned with whether I can remain theory-neutral on the issue of the nature of emotions.

---

[43] In the fourth chapter, I will demonstrate and defend the use of causal evidence in the moral psychology literature to draw concept constitution claims for the emotion theory.

Due to space concerns, I will not be taking a complete stand on what emotions are for the emotion theory of concepts, although certain restrictions will be given and we will discuss what the potential nature and psychological constituents of an emotion may be. While we obviously do not have the space to provide a complete theory of what mental states constitute emotions, I see it as beneficial for the emotion theory that it remains to a significant extent theory neutral concerning taking a stand on what emotions are. In this fashion, philosophers and psychologists can examine and explore the emotion theory of concepts without necessarily having to worry about any meta-foundational issues concerning a theory of emotions. This is somewhat similar to how a philosopher of physics may write a paper on quantum mechanics without having to defend scientific realism in the same paper. Moreover, theorists still can discuss the emotion theory with each other despite the controversial nature of the topic of what emotions are and even if the given theorists disagree on a theory of emotions.

However, there may be an initial issue as to whether the emotion theory is distinct from other theories of concepts in that emotions may themselves be made up of concepts; concepts that may form appraisal judgments. The major concept theories such as the prototype, exemplar, and theory-theory all claim that concepts decompose into further concepts. If emotions also decompose into concepts that form appraisal judgments, then the emotion theory potentially may not be a different theory from the others. In the philosophy of emotions, a *purely* cognitivist view generally claims that emotions are only constituted by appraisal judgments. For instance, Martha Nussbaum, a pure cognitivist,

writes, "[e]motions are appraisals and value judgments."[44] Pure cognitivists claim that all emotions are only constituted by this specific kind of cognitive state. If emotions are just appraisal judgments, then in this instance, the emotion theory may potentially be nothing more than, for example, the prototype theory since concepts that are composed of emotions are really decomposable into further relevant concepts that form appraisal judgments.

However, purely cognitivist views have come under attack from psychological findings. For instance, some such as Lazarus argue that the feeling of emotions is caused by appraisal judgments.[45] Therefore, all emotions are constituted by such judgments. While it is wholly unclear as to how causation can lead to a constitution claim for emotions, even if we grant this move for the sake of argument to the likes of Lazarus, Zajonc has argued that emotions can be induced without any prior cognitive mental states such as appraisal judgments.[46] For example, emotions can be directly evoked through physical means by drugs or hormone treatments without any prior emotion related appraisal judgments occurring in the subject's mind. Moreover, there is support that there is a direct subcortical pathway from the retina to the amygdala, a phylogenetically primitive emotional region of the brain, which bypasses cognition regions that are

---

[44] Martha Nussbaum, *Upheavals of Thought: The Intelligence of Emotions*, (Cambridge: Cambridge University Press), 2001, 4.
[45] R. S. Lazarus, "On the primacy of cognition," *American Psychologist* 39: (1984), 124-129.
[46] Robert Zajonc, "On the primacy of affect," *American Psychologist* 39: (1984), p. 117-123.

associated with appraisal judgments.[47] It is thought that many human beings have an immediate fear response in seeing snake-like coiled objects even before any cognitive appraisal judgments regarding the snake-like object may even be made. Lazarus responds that even though there may be a direct pathway to some phylogenetically primitive subcortical neural structure that correlates with emotions, such a structure may still be able to harbor the capacity to produce appraisal judgments.[48] However, keeping in mind that the subcortical pathway in question bypasses cognitive regions of the brain, the problem with this is that damage to the amygdala can lead to significant deficiencies in affect, but such damage does not impair one's general ability to think about and formulate the cognitivist's appraisal judgments, such as those concerning blame and praise.[49] Therefore, it is doubtful that the amygdala harbors the capacity to produce appraisal judgments. When having, for example, an immediate fear response to snake-like coiled objects in which cognitive regions are bypassed, there are no appraisal judgments involved in the constitution of the emotion. Thus, there is strong empirical evidence against purely cognitivist views of emotions, and I understand such views to be untenable.

On the opposite end of the spectrum, one may claim that all emotions may be non-cognitive and contain no appraisal judgments. One may be a pure non-cognitivist

---

[47] J. E. LeDoux, *The Emotional Brain*, (New York: Simon and Schuster, 1996).
[48] Lazarus, ibid.
[49] See also, M. Siebert, H. Markowitsch, P. Bartel, "Amygdala, affect and cognition: evidence from 10 patients with Urbach–Wiethe disease," *Brain* 126: (2003), 2627-2637.

and hold that all emotions do not contain appraisal judgments. This differs from the position of regular non-cognitivism in that this latter view allows for the possibility that some emotions may contain appraisal judgments, but not all emotions may do so. If pure non-cognitivism is the case, then the emotion theory is clearly distinct from the other concept theories, and the present issue or worry at hand need not concern us.[50]

However, as another potential option, one may hold a hybrid view where some but not all emotions contain appraisal judgments as well as some other mental state component(s) that are essential to all emotions. While there are many potential components of emotions in the literature, I take it that the other three major possible components of an emotion are a felt qualitative "what it is like" aspect, a somatic component where emotions are perceptions of bodily changes, or an action tendencies component where emotions are dispositions to act. Such a hybrid view may be problematic in that emotions at times may once again be constituted by appraisal judgments since emotions in some instances are in part constituted by such judgments. But if such a hybrid view is the case, then at times an emotion is a conjunction of an

---

[50] One objection against pure non-cognitivism is that we make many distinctions between different types of emotions. However, it appears that a pure non-cognitivist theory cannot account for this diversity of distinctions we make based on just feels, perceptions, or action-tendencies. It seems that we must introduce some kind of cognitivism to account for the diversity of distinctions we make for emotions. While there is an ensuing debate about this issue, we need not pursue it here for our purposes given that pure non-cognitivism poses no threat to an emotion-based theory of concepts being different from the prototype, exemplar, and theory-theories. For our purposes, we may assume that pure non-cognitivism is a possible viable candidate for a theory of emotions, especially taking into account the additional fact that in this dissertation, no complete stand on a theory of emotions is taken.

appraisal judgment and some other factor(s) *x*. For our present concern, we need not specify what *x* is, although *x* may be any combination of the above three discussed potential components. Thus, on this theory of emotions, the emotion theory will differ from the other concept theories due to the *x* component, which the concepts that are not composed of emotions will not have. In this respect, emotion theory concepts in some cases may be composed of concepts that form appraisal judgments, but they will also be made up of *x*. Due to this conjunction, emotions on this hybrid view still will be different from prototypes, exemplars, and theory-theories. Here, emotions may in part be constituted by appraisals, but appraisals cannot be constituted by emotions. This difference allows us to escape the worry at hand in that the emotion theory is still a different view of concepts from the others regardless of whether emotions are this specified type of hybrid or whether emotions contain no appraisal judgments at all.[51]

Emotions are constituted by an *x* factor on both of these possible available theories of emotions. Although on the hybrid view, emotions at times may be made up of

---

[51] Another objection one may have to the emotion theory is that concepts can be triggered endogenously, but it appears that emotions cannot be. Concepts can be under the control of and be activated by the agent (endogenous control) rather than be under the control of the environment (exogenous control). Thus, since emotions appear to be only under exogenous control, concepts cannot be constituted by emotions. While much of emotional experience appears to be under exogenous control, it is important to first point out that non-emotion theory conceptual knowledge in the subconscious is also under exogenous control. Second, it seems that we can purposefully excite emotions in our imaginations. For example, when deciding whether I should ride the rollercoaster, in order to make my choice, I may imagine the fear and elation I would feel if I was on the rollercoaster. Also, certain groups such as the stoics and Buddhists are well-known for their ability to control what emotions they do and do not feel. Thus, emotions do appear to be under endogenous control to an extent, and they are under such control especially for those with the proper training.

appraisal judgments as well, we cannot claim that the emotion theory of moral concepts is viable merely based on the evidence for prototype, exemplar, or theory-theory structure. For, what is essential to all emotions is some *x* factor rather than appraisal judgments. Even though it is not specified what this *x* factor is, we may still determine whether emotions influence categorization, which may lead to the viability of the emotion theory. The second objection asks, "How is this possible without knowing exactly what emotions are?" If we closely examine how cognitive scientists determine what emotions are, one method may be seen as a bottom-up approach in which they stimulate what may be potential components of emotions to see if they lead to the experiencing of an emotion in subjects. Whether and what emotions are experienced may be determined by providing a written questionnaire at an appropriate time after the stimulus is provided that asks participants what emotions they may be presently experiencing. For instance, Lazarus and Alfert have run studies in which subjects are given different written descriptions of a graphic film they are about to see. These descriptions influenced what type of emotion was felt by subjects when viewing the film. What emotions subjects felt was determined by providing them the Nowlis Adjective Checklist of Mood that scores for eight variables such as anxiety, pleasantness, and social affection. This study purportedly shows that given the influence of appraisals, some emotions at times may be constituted by them.[52]

---

[52] R. Lazarus and E. Alfert, "Short-circuiting of threat by experimentally altering cognitive appraisal," *Journal of Abnormal and Social Psychology* 69: (1964): 195-205.

Another method may be viewed as a top-down approach in which cognitive scientists induce emotions in subjects, examine subjects that are already experiencing certain emotions, or study patients with blunted emotions in relation to normal subjects and then attempt to discern the *x* factor or what constitutes emotions. For cases of inducing emotions, whether and what emotions are felt by subjects after they are potentially induced by emotions may also be determined by providing a written questionnaire to them. For example, Smith and Lazarus asked participants to imagine themselves being in emotionally charged scenarios such as imagining that there is a great threat and danger that one's relative will die.[53] This particular vignette attempted to elicit fear in the participants, and whether such fear was in fact elicited were determined based on subjects' reports. Participants then answered several questions in respects to how applicable certain appraisal judgments were to a given scenario, where each circumstance attempted to elicit a particular emotion. Given that the above experimenters had adopted a particular cognitivist theory of emotions in which specific appraisal judgments correspond to specific emotions, their theory of emotions predicted that there will be a correspondence between particular appraisal judgments and particular emotions in the experiments. For instance, in the threat vignette that elicits fear, subjects should relate thoughts about physical danger and a low coping potential to the particular situation. In

---

[53] C. A. Smith and R. S. Lazarus, "Appraisal components, core relational themes, and the emotions," *Cognition and Emotion* 7: (1993), pp. 233-269.

this manner, the experimenters attempt to establish some evidence that emotions are at least in part composed of appraisal judgments.

Given the availability of the top-down approach in cognitive science, we can determine whether an emotion influences moral judgment without specifying the *x* factor that constitutes the emotion. Simply put, we may induce emotions in subjects by, for example, asking them to recall an emotional moment of their lives or deal with patients who are already known to have deficiencies in affect, without knowing what essentially constitutes the emotion. At this point, we may then attempt to discover what the *x* factor is if we so choose; a project that will not concern us in the dissertation. At this same point, we may also then run other tests to see whether the induced emotion influences reasoning or see whether the fact that the blunted emotions of certain subjects is experimentally responsible in some way for being able to infer the influence of emotions in the higher competences in normal subjects, which may lead to the viability of the emotion theory; a project we will be very interested in in the dissertation. This general strategy will be adopted in attempting to prove the viability of the emotion theory later in this essay while simultaneously not providing a complete account of what emotions really are.

Another objection may be that on the emotion theory, concepts are constituted by emotions. Concepts have intentionality and content in that they are mental representations that are about, of, or represent properties in the world. Given that some concepts may be composed of emotions, these emotions must also be mental

representations. Yet, how can this be the case if you leave open the possibility that emotions at times may not be composed of appraisal judgments? While it is obviously beyond the scope of this essay to offer a theory of intentionality for emotions, how can emotions represent if emotions are not composed of concepts that are the constituents of appraisal judgments; appraisal judgments which represent propositions? The strategy here is to lessen the worry of this second criticism by offering a plausible way for emotions to be intentional mental states even if emotions may be non-cognitive.

In the emotions literature, I generally take it that many to most believe emotions represent material and/or formal objects. The material object can be understood as a specific object while the formal object is a description of the material object that must hold for the material object in order for the emotion in question to be appropriately instantiated. For example, the feeling of fear when being chased by a lion is materially directed upon the lion and is formally about danger, where the lion is represented as being dangerous. My emotion of anger may be materially about my enemy but formally be about the fact that an offense has been made against me. This is generally thought to differ from moods in that my general mood of being anxious may not be directed upon any object or state of affairs. Now, some theorists may believe that emotions can only refer by virtue of the fact that emotions are composed of appraisal judgments. As appraisal judgments are about propositions or states of affairs, there is no dilemma in claiming that emotions are mental representations so long as emotions are composed of a string of concepts that form an appraisal judgment. Some cognitivist oriented

philosophers have taken this as an indication that emotions must be constituted by appraisals since emotions are intentional mental states. However, since I leave open the possibility that some emotions may not be composed of appraisals, how can such emotions be mental representations?

Even though some emotions may not be composed of appraisal judgments, we may potentially lessen the urgency of this criticism in that many naturalistic theories of content such as an informational or even a number of teleosemantic views allow for a mental state to represent even though what constitutes the mental state does not itself describe its content. In other words, the complexity of a mental state's content need not be mirrored within the structure of the mental state itself. There are numerous variations and refinements of informational and teleosemantic theories. Informational theories may be understood to claim roughly that mental representation R represents C so long as R is reliably caused by C. A teleosemantic theory generally might instead claim, for example, that R represents C only if R has been selected for the job of indicating C. While no stand will be taken on a particular naturalistic theory of content, such widely discussed theories generally allow for emotions to represent even though the emotion in question may not be constituted by an appraisal judgment. Due to obvious space concerns, even though I do not argue that any of these relevant naturalistic theories of content are correct, I attempt to lessen the immediate worry of this second objection by showing that many popular naturalistic views of content, if any of them may be correct, can account for the concern at hand.

Although I do not necessarily espouse his view of intentionality for emotions, Prinz's theory provides a nice illustration of how emotions may plausibly represent even though they are not themselves constituted by appraisals.[54] Prinz maintains a purely non-cognitivist view of emotions in which emotions are not composed of appraisal judgments. However, in line with Fred Dretske's later and general theory of mental content, Prinz offers a hybrid informational and teleosemantic theory of intentionality for emotions.[55] He claims that emotions represent formal objects that are Lazarus's core relational themes, where core relational themes are formal relational properties that may pertain to one's well-being, and such formal relational properties reliably cause one to have certain emotions. For example, anger can represent the core relational theme offense, and fright is about the core relational theme concrete dangers. An offense reliably causes one to feel anger while concrete dangers reliably cause one to feel fear.

Moreover, the teleosemantic element enters the equation when he claims that emotions have the function of being caused by core relational themes in that there is an evolutionary advantage to have emotions that detect such themes. For instance, anger in part motivates us to respond when material goods have been taken away from us, and fear confers a survival advantage in that it protects us from dangers. On this hybrid theory of intentionality for emotions, emotions can represent core relational themes even though

---

[54] Jesse Prinz, *Gut Reactions*, (Oxford: Oxford University Press, 2004).
[55] Fred Dretske, *Knowledge and the Flow of Information*, (Cambridge, MA: The MIT Press, 1981). Fred Dretske, "Misrepresentation," in *Belief: Form, Content and Function*, ed. By R. Bogdan, (Oxford: Oxford University Press, 1986), pp. 17-36.

emotions are non-cognitive. Similar to how the beeping of a police detector in one's car represents the presence of a police radar but the beeping itself lacks structure and cannot be semantically split into a tone meaning "police" and another meaning "radar," emotions can represent so long as the emotion stands in the proper mind-world causal relation to C, and the emotion has the function of indicating C even though the emotion itself is not constituted by appraisal judgments.

Moreover, a concept of *x* is thought to be constituted by mental representations that are about *x* or about a property of *x*. For example, my LION concept, which represents the category *lion*, may be in part constituted by FOUR LEGS, MANE, YELLOW, and ROARS, where, for example, YELLOW refers to or represents the property of *being yellow*; a property that lions have. While the answer to whether an actively rendered non-cognitive emotion that constitutes a moral concept relevantly can be about *x* or about a property of *x* is in part dependent on what particular theory of intentionality is espoused for emotions, it is quite plausible that a positive answer may be given to this question. For example, on Prinz's particular view only in regards to intentionality, take the moral emotion of anger that may represent *being an offense*. If my moral concept MORALLY WRONG in my moral judgment is in part made up of the actual and actively rendered emotion of anger, then since anger represents *being an offense* and *being an offense* is normally thought to be a property of the category *morally wrong*, MORALLY WRONG may be constituted by the

emotion of anger in that anger is about a property of the category *morally wrong*.[56]

[56] While this example clearly works if moral properties are mind-independent, if moral properties are mind-dependent where specifically moral properties are themselves constituted by emotions, then there may be some confusion. For, if a moral property is constituted by anger and anger refers to *being an offense*, where *being an offense* may be thought to be a property of the moral property, then how can anger reliably detect a property that is a property of itself? While this may be a salient objection against views that claim that moral properties are constituted by emotions, no stand is explicitly taken in this dissertation concerning the nature of moral properties. This dissertation is primarily about the nature of moral concepts rather than of moral properties. If this potential objection does indeed work, then we may make a general underlying assumption here that moral properties cannot be constituted by emotions.

3. The Classical, Prototype, and Exemplar Theories

This chapter will assess the viability of the classical, prototype, and exemplar theories for moral concepts.

3.1 The Arguments Against the Classical Theory

There exist several general arguments against the classical view for non-mathematical and non-logical concepts. A common objection is that for over two thousand years, philosophers have attempted to arrive upon definitions for concepts such as JUSTICE and KNOWLEDGE with no success.[1] There is still no consensus on such definitions, and if there is a definition for them, it should have been discovered by now. Given that such definitions have not been discovered, this provides reason to believe that many of our concepts generally do not have classical structure. For, if our concept of, for instance, JUSTICE might embody definitional structured knowledge, then it is reasonable to expect that we should have come upon it by now. The fact that for well over two thousand years we have not arrived upon knowledge of necessary and sufficient conditions for JUSTICE lends some support to the inference that perhaps our JUSTICE concept does not have classical structure.

---

[1] By "KNOWLEDGE," I mean the concept that is traditionally defined in philosophy as true justified belief rather than as an information carrying mental state.

Another objection is made by Wittgenstein in his analysis of GAME.[2] He claims that the features used to classify things as games are not necessary for membership. For instance, *involving the use of a ball* is not a feature because chess and poker are games that do not include the use of a ball. *Involving two or more people* is not a necessary characteristic due to games such as solitaire that are played by oneself. Rather than having a definition, Wittgenstein suggests that we classify things as games based on his aforementioned family resemblance, which, as previously stated, is the main philosophical influence for the prototype view.

Another objection against the classical view is related to the typicality effects in prototype studies. While psychologists such as Piaget have assumed definitionism to be true, typicality effects have put this theory in serious jeopardy. In positing that we have in mind necessary and sufficient features when representing a category, the classical view predicts that the members of a category should be considered as equal tokens, where no member is considered to be a better or more typical instance of the class. After all, in order to be a member, all must equally satisfy the strict definition. However, as previously mentioned, cognitive scientists have discovered that for concrete concepts, there are such typicality effects or graded membership, where members do stand in a graded relationship with each other. This is not predicted by definitionism, and such effects actually stand as contrary to the equal membership definitionism predicts. Moreover, the inability of definitionism to explain typicality effects also includes its

---

[2] L. Wittgenstein, *Philosophical*.

inability to explain how typical members are categorized more quickly than atypical ones. While views such as the prototype, exemplar, and theory-theory can provide an explanation for such phenomena, the classical theory cannot.

Second, researchers have discovered through prototype studies that human beings do not use definitions in the higher cognitive competences, at least for non-mathematical and non-logical concepts. For instance, prototype studies have shown that subjects frequently list features of members of a category that are not necessary and sufficient qualities. For example, members of the *bird* class may have the features *flies*, *sings*, and *has wings*, but these qualities are not individually necessary and jointly sufficient for an individual to classify an object under *bird*. An object may still be classified as a bird even though it has lost its wings and does not fly or sing. Given that the features attributed to members of a class are not necessary and sufficient conditions and that there is no evidence for definitionism, cognitive science has shown that humans generally do not use definitions in the higher cognitive competences.

The above psychology-based arguments lead to the generally held conclusion that definitions most likely are not psychologically real. Now, some may claim that humans psychologically still do have definitional knowledge as well as some other body of knowledge such as prototypes. Recall that several of the early hybrid theories such as Osherson and Smith's are set up to make this claim. Remember that for them, prototype features are used in fast categorization, but when one must really classify things correctly, definitions are used. However, from a psychological standpoint, it has been

shown that the classical view does not partake in the scientific explanation of concept phenomena such as categorization and typicality effects. Moreover, psychology studies have been successfully conducted in which opposite predictions from that of definitionism are born out. The claim that a hybrid theory that contains definitionism may be correct and that we still have definitional knowledge is empty in that there currently is no evidence for the psychological reality of classical structure, and it plays no role in the scientific explanation and prediction of psychological data. Now, some philosophers such as Rey escape this psychological reality objection by claiming that the classical view is in regards to correct concepts rather than conceptions. However, as mentioned, our primary concern here is only in regard to conception theories, and such a move by Rey need not be addressed.

If the classical view is not viable for moral concepts due to such above related reasons, some may argue that through further reflection upon our moral beliefs we psychologically may be able to arrive upon components that represent necessary and sufficient conditions. While this remains a possibility, as it also does with concrete concepts, given the numerous prototype studies to be discussed below, there is no scientific evidence consistent with PAC to justify the expectation that this may be so. As it will be shown that the classical view currently plays no role in scientific explanation and prediction in our moral psychologies, even though it may be possible that future studies will demonstrate the viability of definitionism, a level of doubt may be established against this theory such that we are warranted through inference to the best

explanation in concluding that a classical structure for moral concepts is most likely not psychologically real. This conclusion is further supported by the previously mentioned fact that for over two thousand years, we have not yet arrived upon a set of necessary and sufficient conditions for moral terms.

In the following section, we will see positive evidence for moral concepts that, like concrete concepts, produce typicality effects and that show that subjects do not use or rely on definitions for categorization. Such evidence, as previously indicated, will be taken to provide the ammunition for making arguments against definitionism for moral concepts.

3.2 The Arguments for Prototype Theory

A few philosophers have adopted the prototype theory as a viable view for moral concepts. For instance, Paul Churchland, in *A Neurocomputational Perspective*, can be read as providing an account of moral concept acquisition by positing the existence of moral prototypes.[3] Churchland argues that learning about morality is more about learning *how* rather than a matter of learning *that*. In other words, moral concept acquisition is about learning how to recognize various complex situations and how to appropriately respond to them by way of an at times long and painful process of social learning rather than by applying abstract moral principles. On this view, certain moral

---

[3] Paul Churchland, *A Neurocomputational Perspective*, (Cambridge, MA: The MIT Press, 1992), 297-303.

situations will activate prototypes given the situation's similarity to one's previous experiences and training in other moral situations. Through social experience and further development, a child's moral prototype and practical wisdom can change and expand by being impacted by factors such as how society expects them to react to certain situations. For instance, a child may be guided in life primarily by self-interest. However, later the child comes to realize that her society generally expects her to take into consideration others' feelings and interests in certain moral scenarios. Fully taking into account other people's considerations leads to a change in her moral prototypes.

While the enthusiasm and insight of applying the prototype theory to moral concepts is duly noted, Churchland does not support the fact that moral concepts have prototype structure with the appropriate evidence that is required from PAC. Remember that PAC requires as a matter of precaution that all abstract concepts must have experimental support specifically in regards to the particular abstract concept in question in order to make any claims about its conceptual structure. For example, no experimental studies on moral concepts are cited in Churchland's work to warrant the claim that moral concepts have prototype structure. Thus, his prototype conclusion in-and-of-itself is unjustified based on PAC.

While most experimental work on prototype theory has generally focused on concrete concepts, there are a handful of studies on certain concepts that do not appear to be straightforwardly moral concepts but may be somehow related to them. For example, Linda Coleman and Paul Kay have run a study that illustrates prototype structure for the

concept LIE.[4] The issue with LIE is that it appears to be a concept that has components that are in some way or another merely about descriptive acts. Therefore, LIE may not itself be a normative concept. There does not appear to be any explicit rightness or wrongness associated with the concept LIE, as opposed to if LIE is a constituent in the judgment, ONE OUGHT NOT LIE, where this entire thought itself is about a normative principle. It will be important to keep this potential problem in mind while we examine Coleman and Kay's studies. Now, this initial test is replicated and supported by two further experiments on the concept by Eve Sweetser on the one hand and Abigail Strichartz and Roger Burton on the other.[5] In the Coleman and Kay experiment, they ran a study to see if the following features that are represented by the constituent components of LIE are defining necessary and sufficient conditions or whether they are a summary representation of features for folk in determining when a particular act is a lie:

1) P is false.
2) S believes P to be false.
3) In uttering P, S intends to deceive A.[6]

Thus, in essence, they tested for falsehood, deliberate falsehood, and intent to deceive. They provided participants with a host of situations that contained eight various combinations of the three given features; features that may be present or absent in the

---

[4]Linda Coleman and Paul Kay, "Prototype Semantics: The English Word Lie," *Language* 57:1 (March 1981), 26-44.

[5] Eve Sweetser, "The definition of lie: An examination of the folk models underlying a semantic prototype," in *Cultural Models in Language and Thought*, ed. D. Holland and N. Quinn (Cambridge: Cambridge University Press, 1989), 43-66. Abigail Strichartz and Roger Burton, "Lies and Truth: A study of the Development of the Concept," *Child Development* 61:1: (Feb., 1990), 211-20.

[6]Coleman and Kay, Ibid.

various combinations, where subjects were asked to circle whether the situation is a lie, is not a lie, or unknown. They were also questioned on the typicality of each situation as an instance of lying or not lying. Coleman and Kay found that the more of the three features of lying that a situation contained, the more typical or better instance of a lie the situation is deemed to be. Moreover, they discovered weighted features, where 2) is the most heavily weighted and 3) is next in importance. Strichartz and Burton, who were able to successfully replicate this study, also discovered that the weight placed on each of the features differs based on age or maturity.[7] Also, situations that contained any variation of two of the features tended to have a mean score where such situations are still considered as lies. Typicality effects, weighted features, and the fact that the three given features are not held to be necessary and sufficient conditions for determining category membership leads to the conclusion that LIE has prototype rather than definitional structure.

Another study in which the concept in question may or may not be a moral concept is Hampton's aforementioned study on abstract concepts that examined the concept CRIME.[8] This too, like LIE, initially appears to have components that are only about a descriptive kind of act. He was able to produce typicality effects and reach the conclusion that CRIME has prototype structure. By way of a feature listing task where participants list what features they think belong to a category, he was able to discover that participants associated features such as *an act*, *adverse effect on victim*, and *is done*

---

[7]Stichartz and Burton, Ibid.

[8] Hampton, "An Investigation of the Nature of Abstract Concepts."

deliberately with the abstract category *crime*, where such features have different weights and account for typicality effects. Furthermore, such features are not necessary and sufficient conditions for determining category membership based on the fact that some situations are deemed to be crimes, but they do not satisfy all of the perceived features of crime.

The question now is whether the perceivably descriptive prototype concepts LIE and CRIME are in fact normative moral concepts that have prototype structure. For, if they are, then this shows that in some cases prototype theory is a viable theory of moral concepts. Mark Johnson argues that specifically LIE is in fact a moral concept that has prototype structure.[9] He makes this claim based on Sweetser's insight that the prototype structure of LIE is dependent on one's background knowledge or what George Lakoff's calls an idealized cognitive model.[10] As Lakoff contends, "category structures and prototype effects are by-products of [idealized cognitive models]."[11] Idealized cognitive models are structural organizations of our background knowledge, where such models are themselves exercised in the higher cognitive competences. They are similar to a complex version of the theory-theory of concepts in that they associate general background knowledge and beliefs with a concept. Idealized cognitive models or organized abstracted background beliefs and information we have acquired through our experience with the world becomes selected during cognition based on the context or situation we

---

[9] Mark Johnson, *Moral Imagination*, (Chicago: The University of Chicago Press, 1993), 95-98.
[10]George Lakoff, *Women, Fire, and Dangerous Things*, (Chicago: University of Chicago Press, 1987).
[11]Ibid., 68.

find ourselves in. Such knowledge may be in the form of imaginative structures that help us make order and sense out of the world by aiding us in such things as categorization and concept combination. Moreover, such knowledge provides a way for us to evaluate certain circumstances and judgments based on our previous social experiences. Lakoff claims that idealized cognitive models underwrite the prototype structure of concepts.

Johnson claims that LIE is dependent on the idealized cognitive model of "ordinary communication." This cognitive model may account for the fact that communication is even possible. Thus, such a cognitive model provides the requisite background information when we converse with others, such that we have the proper expectations that in normal conversations we may assume such things as that the truth will normally be told among human beings given that we do not always have to worry that any and every person we talk to could very well be lying to us. Thus, we have:

> *The Idealized Cognitive Model of "Ordinary Communication"*
> (1) People intend to help rather than harm one another.
> (2) Truthful information is helpful.
> (3) The speaker intends to help the hearer by sharing information.
> (4) A Speaker who knowingly communicates false information intends to harm the hearer.[12]

Since the cognitive model of "ordinary communication" accounts for the background beliefs (1)-(4) that most people have with respects to *lie* and the prototype effects for LIE are grounded in this cognitive model as well, Johnson claims that LIE is a moral normative concept.

---

[12]Lakoff, *Women, Fire, and Dangerous Things*.

> [W]e need only remind ourselves why lying is usually considered to be a bad thing to do. The answer is that lying is typically *harmful* to others. But notice that the relevant notion of harm is partially specified by the idealized cognitive model of ORDINARY COMMUNICATION. A notion of helping and harming is *built into* that model. Therefore, to the extent that our determination of what counts as a lie in a particular situation depends on the idealized cognitive model of ORDINARY COMMUNICATION, our understanding of *lie* is inextricably tied up with an evaluative notion of help and harm.[13]

Johnson may be on the right track here in that he appears to want to bring in the notion of the moral wrongness of harming others as part of "ordinary communication." Insofar as there is some connection between the prototype of LIE and the idealized cognitive model, LIE may then be considered to be a normative moral concept that has prototype structure. However, the first problem is that if we closely examine the four beliefs in the idealized cognitive model for "ordinary communication," we still do not see anything explicitly normative. The concepts HELPING and HARMING, by themselves, may be thought of as merely being about a descriptive class of acts that have to do with alleviating or instigating physical or psychological harm from or to another human being, respectively. Therefore, since there is no explicit notion of moral wrongness in (1)-(4), it still may not be the case that LIE is a moral concept.

Second and what is even more problematic is that since the idealized cognitive model is claimed to be one of the fundamental structures of the abstract concept LIE, no experimental evidence consistent with PAC is given to indicate that LIE has such a structure. Idealized cognitive models such as "ordinary communication" are purported

---

[13] Johnson, 95-96.

structural aspects of LIE, but no empirical data on this potential moral concept has been collected to warrant this assertion. Rather, such a conclusion is blindly given based on concrete concept studies and speculation, which is in violation of PAC. Therefore, insofar as idealized cognitive models underwrite the prototype structure of moral concepts and supposedly allows LIE to be a moral concept, Johnson's claim that LIE is a moral concept does not satisfy PAC, and he does not have the appropriate evidence to show that it is a moral concept.

The question still remains as to whether LIE may be shown to be a moral concept in a manner where the concept has prototype structure. In this way, we may conclude that the prototype theory is a viable theory for some moral concepts. In order to accomplish this, we will need to bring in the aid of Bernard Williams and his discussion of thin and thick moral concepts; a view which is influenced by John McDowell.[14] Williams argues that there are thin moral concepts which appear to be purely normative such as RIGHT, GOOD, BAD, and OUGHT. On the other hand, there exist thick concepts that have both a normative and descriptive component to them; components that are not separable and are culturally formed. Williams does explicitly list LIE as a thick concept among others such as PROMISE, BRUTALITY, COURAGE, and GRATITUDE.[15] For example, LIE has a descriptive or factual component in that such a component is typically about a

---

[14]Bernard, Williams,*Ethics and the Limits of Philosophy*, (Cambridge, MA: Harvard University Press), 1985. John McDowell, "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society*, suppl. Vol. 52: (1978); "Virtue and Reason," *Monist* 62: (1979).
[15]Bernard Williams, *Ethics and the Limits of Philosophy*, 140.

particular kind of verbal act of deception and members of a particular community that have this concept can point out when an instance of lying has occurred. However, it also has a normative or value component that is about "oughtness" which provides reasons for action that generally one ought not to lie. It is contended that these features the two components represent are necessarily intertwined and inseparable in that they are represented by the components of culturally-conditioned concepts, where, as McDowell argues, an outside observer of a community cannot fully pick up and understand the features represented by the descriptive component of a community's thick moral concepts without also imaginatively grasping the features represented by the normative evaluative component of the concept. Thus, Williams and McDowell argue that the fact/value distinction has not been found or discovered in the concepts and language of ethics, but rather, it has been artificially brought there by the minds of philosophers such as R. M. Hare.[16] What we can see here is that even though *lie* first appears to have features pertaining to descriptive acts, it generally is thought to have a negative normative feature attached to it in the United States, where there is an initial defeasible presumption that acts of lying are wrong. Something similar to this may be at work for *promise* as well, albeit in a positive normative sense. It is this normative aspect that is represented by the components of LIE that allows LIE to be a proper thick moral concept.

Given that it has been determined that LIE has a prototype structure and that it may be a thick moral concept where members of a community that acquire the concept

---

[16]R.M. Hare, *Moral Thinking,* (Oxford: Oxford University Press, 1982).

understand it to have components that are about descriptive and normative features, we should expect that experimentally, at least one of the components of the prototype for LIE will be about a moral normative feature. However, the three prototype studies for LIE are not feature listing tasks where participants list what features they believe belong to the category a concept represents. Rather, the possible features for the class are already preconceived by the experimenters and given to the participants with no room for expansion. The predetermined features for *lie* are not explicitly normative features. Thus, we cannot know for certain whether participants attributed a normative moral feature to the class. However, Hampton's feature listing task study for *crime*, which was run in the U.S., did show that the corresponding concept had prototype structure in which one of the features for *crime* frequently listed by participants appears to be morally normative: *breaks the moral and social code.* While MORAL is typically understood to be a normative concept which would mean that CRIME must itself then be a normative concept since it is in part constituted by MORAL, one may object that perhaps BREAKS THE MORAL CODE is being understood in an anthropological sense where it is merely a neutral descriptive thought had by participants of what a culture thinks about a crime. While this is a possibility, we will later examine a study on HIGHLY MORAL PERSON. While this compound concept also contains the concept MORAL albeit in a different context as compared to the case we are presently examining, what constitutes this compound concept may help us to inductively infer whether the folk generally understand BREAKS THE MORAL CODE to be a normative rather than an anthropological component

of CRIME. The study on HIGHLY MORAL PERSON found that most participants' concepts contained what are usually thought to be normative virtue concepts such as HAS INTEGRITY and HONORABLE. It also contained concepts such as EXEMPLARY and RESPECTED that imply that one ought to imitate highly moral individuals because they are praiseworthy. Moreover, if subjects viewed morality to be anthropologically descriptive rather than prescriptive, then we should expect their components of HIGHLY MORAL PERSON to contain complex concepts such as FOR OUR CULTURE, FOR THIS SOCIETY, or something to this effect that qualifies the components this concept may have. However, no such anthropological qualifiers were found. Based on the fact that most participants' HIGHLY MORAL PERSON concept is constituted by what are typically thought to be normative concepts, concepts of or related to ought-to-be-pursuedness, and that there are no relevant anthropological qualifications to the components of the compound, it appears that a reasonable inductive inference may be made that when simply moving from thoughts concerning MORAL PERSONS to MORAL CODES, the thought BREAKS THE MORAL CODE may likely be understood to also have like normative components *mutatis mutandis* rather than anthropologically descriptive ones. Naturally, a stronger inference may be made if the same subjects used in Hampton's CRIME study were also used in the HIGHLY MORAL PERSON experiment or if Hampton's study directly questioned the nature of BREAKS THE MORAL CODE. While such studies have not been conducted, given the present circumstances, a moderate inference still may be made that allows us to conclude

that it is likely that participants in Hampton's study for CRIME think of BREAKS THE MORAL CODE in a normative rather than anthropological light.

Therefore, in this fashion, since CRIME may now be shown to be a thick moral concept,[17] Williams' contention for dual components to thick concepts is to a certain extent empirically vindicated. Given that CRIME is a thick culturally-formed moral concept with a component that represents a moral normative feature and that it has prototype structure, we may conclude that some moral concepts have prototype structure, although it has not been empirically proven nor need it be proven for our purposes that the dual components are inextricably linked. All that needs to be shown is that there is a normative component to CRIME which then qualifies it as a thick moral concept in the U.S. and most likely in other cultures as well that view crimes in a negative normative light. Nothing need be said here about whether the normative component is in principle separable from the descriptive component. Furthermore, a predictive inference may be drawn that if a feature listing task for *lie* is actually conducted, a statistically frequent moral normative feature will most likely be given by subjects since it is the case that such a feature has been discovered to exist for *crime*.

Also, notice that the given conclusion that the prototype theory is a viable theory of moral concepts has remained true to PAC. By using experimental data on certain specific abstract thick concepts that initially appear to refer to only a descriptive class of

---

[17] As will be clarified in the fourth chapter, that moral concepts have prototype and exemplar structure is a prima facie claim.

acts along with the application of philosophical argument regarding such concepts, we have discovered that some moral concepts have prototype structure.

At this juncture, I will introduce several experiments on specific concepts that are more clearly within the moral domain. Jennifer Frei and Phillip Shaver have run a study on the abstract concept RESPECTFUL TO A PARTNER, which may be considered to be more directly a moral concept in that it is a compound virtue concept.[18] This study found that this compound concept is constituted by such things as SENSITIVE TO FEELINGS, FOSTERS EQUALITY, CARING, and LISTENS TO MY VIEWPOINT. Typicality effects were discovered for this complex concept. Furthermore, the list of features was not taken by participants to be necessary and sufficient conditions for determining whether a potential member belonged to the category. Hence, insofar as RESPECTFUL TO A PARTNER is a moral virtue concept with prototype structure, this provides further evidence that prototype theory is a viable theory for moral concepts.[19]

---

[18] Jennifer Frei and Phillip Shaver, "Respect in close relationships: Prototype definition, self-report assessment, and initial correlates," *Personal Relationships* 9: (2002), 121-39.

[19] On an aside, since virtues have been found to have prototype structure, such a structure may be able to explain why Aristotle believes that some mean virtues appear to lean closer to one extreme rather than another. In providing the doctrine of the mean in the *Nichomachean Ethics*, Aristotle claims that virtues are means that lie between two vices of deficiency and excess. For example, courage lies between the two extremes of cowardice and rashness. However, some virtues appear to be closer to one extreme rather than the other. For instance, courage appears to be closer to rashness than cowardice. "E.g., since rashness is thought liker and nearer to courage, and cowardice more unlike, we oppose rather the latter to courage; for things that are further from the intermediate are thought more contrary to it." Aristotle contends that the mean is difficult to hit, so at times we should aim for the excess vice that appears closer to the mean. Prototype structure for virtues may explain why one extreme for a particular virtue appears closer to it rather than the other extreme because the closer extreme shares more prototypical and

Also, in a separate study, Lawrence Walker and Karl Hennig, in a feature-listing task, discovered prototype structure for compound concepts that contained virtues.[20] Specifically they discovered prototype structure for JUST PERSON, BRAVE PERSON, and CARING PERSON. For instance, for JUST PERSON, they found that some of the statistically frequent components were LISTENS TO ALL SIDES, FAIR, MORAL, and TRUTHFUL. For BRAVE PERSON, some of the statistically frequent components were HEROIC, FACES DANGER, STANDS UP FOR BELIEFS, and GALLANT. For CARING PERSON, some of those components that passed the production frequency measure to be summary representation components were SYMPATHETIC, GOOD-HEARTED, LOVING, and NURTURING. Common to many prototype theories, for all three compound concepts, they discovered that some of the components were weighted as more important than others. Thus, a potential member will be categorized in a class if she matches the features represented by the heavily weighted components even though she does not match the features represented by the lower weighted components. Given that components of the compounds are weighted and do not represent necessary and sufficient features, this lends evidence to the conclusion that such moral concepts have prototype structure.

---

more heavily weighted features and is more similar to the mean as compared to the opposing extreme. Whether this is or is not the case will be left for others to decide and examine.

[20] Lawrence Walker and Karl Hennig, "Differing Conceptions of Moral Exemplarity: Just, Brave, and Caring," *Journal of Personality and Social Psychology* 86: (2004), 629-647.

Lawrence Walker and Russell Pitts were able to discover prototype structure for a complex concept that contains a thin moral concept.[21] The complex concept in question is the aforementioned HIGHLY MORAL PERSON. They asked participants to freely list any attributes members of such a category may have. Through a production frequency measure, they found that some of the statistically frequent features subjects listed were: *has clear values*, *self-disciplined*, *principled*, *exemplary*, *maintains high standards*, and *has integrity*. Moreover, they also discovered that of the statistically frequent qualities, certain ones were deemed to be more heavily weighted than others. Moreover, the qualities were not taken to be necessary and sufficient conditions in categorization. Given that the statistically frequent features are weighted and are not taken to be necessary and sufficient conditions, this provides evidence that HIGHLY MORAL PERSON has prototype structure.

Finally, Kyle Smith, et al., in a cross-cultural study, have determined that there is prototype structure for GOOD PERSON.[22] This feature-listing task experiment spanned seven different cultures including Chamorro, Filipino, Taiwanese, Turkish, U.S., Venezuelan, and Palauan participants in which weighted statistically frequent features that are represented by prototypes and that are not necessary and sufficient conditions were discovered in all of the cultures. Here, our experimenters claim that cultures varied

---

[21] Lawrence Walker and Russell Pitts, "Naturalistic Conceptions of Moral Maturity," *Developmental Psychology* 34: (1998), 403-418.

[22] K. Smith, S. Smith, and J. Christopher, "What Defines the Good Person? Cross-Cultural Comparisons of Experts' Models With Lay Prototypes," *Journal of Cross-Cultural Psychology* 38: (2007), 333-360.

substantially in terms of which features they have in mind when representing GOOD PERSON. For example, a couple of heavily weighted components generally held for people in the U. S. are CARING and HONEST. These components were not contained within the set of heavily weighted components for the GOOD PERSON concept for the Taiwanese. Rather, their respective concept was constituted by heavily weighted components such as INDEPENDENT and TOLERANT; components not contained within the set of heavily weighted components for the GOOD PERSON concept for people from the U.S. Given all of the above studies that in part contained concepts such as GOOD, MORAL, JUST, CRIME, and BRAVE, we now may inductively infer that moral concepts generally do have prototype structure.

The viability of the prototype theory for moral concepts sits well with concept acquisition, which is the capacity to acquire concepts. The concept acquisition and categorization desiderata can be thought of as two sides of the same psychological phenomenon – our disposition to put individual items into classes. If we have acquired a concept, then we can categorize items as members of the extension of the concept. Likewise, if we can use a concept to categorize, then at some point we have acquired the concept with its concomitant bodies of knowledge. We have already discussed numerous prototype categorization studies that show that various moral concepts are constituted by prototypes. Such studies may also be used to demonstrate that we acquire prototype knowledge at some point in moral education. The previously discussed prototype categorization studies show that we acquire prototype bodies of knowledge.

Also, one may wonder whether the rules and reasons for action we acquire are necessary and sufficient conditions or whether they are represented by prototypes. However, as previously discussed in this chapter, we have found that there are pervasive typicality effects for moral concepts. As previously shown, such typicality effects along with the fact that for well over two thousand years we have yet to reach any consensus on the definition of moral terms allows us to infer that the classical view is not psychologically real and that the knowledge acquired is prototype rather than definitional bodies of knowledge.

As mentioned in the previous chapter regarding the illustration of the prototype view, with the acquisition of things such as rules, one's prototype BEING GOOD concept through moral education may be constituted by DO NOT MAKE OTHERS FEEL BAD, DO NOT HIT OTHERS, and DO NOT LIE. Here, on the prototype view, components will be weighted where there will be graded membership. For example, DO NOT HIT OTHERS may weigh more in terms of importance for a child than DO NOT LIE because, for example, the child receives more punishment for hitting others than lying or the child's parents and teachers more strongly emphasize that hitting others is bad. Therefore, in this case, the child will judge acts of striking others as more seriously bad than acts of being dishonest.

Also, the components will not represent necessary and sufficient conditions. For example, while there are various formulations of normative ethical theories, a normative theorist may claim that a necessary condition for categorizing an act as being good may

be that the act must match one of the features represented by the components of one's BEING GOOD concept. Therefore, when a child witnesses an act of another classmate not retaliating against a playmate who hit him, the child may categorize it as an instance of being good since the act matches the features *do not hit others* and *do not make others feel bad*. This satisfies the necessary condition. However, *ex hypothesi*, these features are not taken to be sufficient conditions in that the child knows that in certain circumstances it is permissible to hit others, such as when being kidnapped by a stranger. To note, this example is meant to be just a reminder to the reader from the second chapter of how prototype structured concepts may look like and operate. As previously discussed, the primary refutation in this chapter of any kind of classical structured view of moral concepts does not come by way of analyzing such a hypothetical case but in part by way of experimentally-based typicality effects, such as those found by Smith's aforementioned prototype study for GOOD PERSON that was in part run on adolescents and by Strichartz & Burton's previously discussed prototype study for the thick concept LIE that was conducted on nursery school, preschool, and first and fifth graders. It is from such studies as well as a plethora of other prototype experiments that have been previously discussed in this chapter that allow us to infer that children and adults do not have classically structured moral concepts.

3.3 The Prototype-Exemplar Chain Argument for the Exemplar Theory

A host of philosophers have claimed that moral concepts have exemplar structure. For example, Goldman, Stich, and Clark have done so along with suggesting that ethical concepts also may have prototype structure.[23] Wong has gone a little further by explicitly claiming that there is a pluralism between exemplar and prototype structures for moral concepts.[24] Prinz, in *Furnishing the Mind*, has also claimed that there is exemplar structure for ethical concepts.[25] However, the problem with the above thinkers is that none have provided evidence consistent with PAC that exemplar theory is a viable theory for moral concepts, let alone have the select authors who also have adopted the prototype view given such relevant data for it.[26]

There are no psychology studies explicitly showing exemplar structure for moral concepts. Thus, initially it appears that there is doubt as to whether it may be shown that exemplar theory is a viable theory of moral concepts in a manner that is consistent with PAC. Nevertheless, it may still be shown that exemplar theory is a viable theory of moral concepts without exemplar studies. In order to accomplish this feat, we will need to introduce a style of argument in the concepts literature that I call *chain arguments*. Chain

---

[23] Stephen Stich, "Moral Philosophy and Mental Representation," *The Origin of Values*, Ed. by Hechter, M., Nadel, L., and Michod, R. New York: Aldine de Gruyer, 1993, pp. 215-228.Alvin Goldman, "Ethics and Cognitive Science," *Ethics* Vol. 103 (Jan 1993), 337-360. Andy Clark, "Connectionism, Moral Cognition, and Collaborative Problem Solving," in *Mind and Morals*, ed. by L. May, M. Friedman, and A. Clark, (Cambridge, MA: The MIT Press, 1996).
[24]Wong, ibid.
[25]Prinz, *Furnishing the Mind*.
[26] Clark supports his prototype contention by citing Johnson, but as we have seen, Johnson's prototype argument is problematic on several accounts.

arguments are claims that if one structure of concepts is already established as being viable, then based on this fact, a different concept structure may be established as being viable. It is called a chain argument because the established concept structure is linked or chained to another concept structure such that this latter concept structure represents a viable theory of concepts based on the viable existence of the first established concept structure. The argument at hand is the *prototype-exemplar chain argument*.

This argument begins by acknowledging that as previously stated, a prototype is an abstracted *summary* representation of features for a category, and it has been shown in the previous section that through empirical tests and philosophical reasoning, some moral concepts do have prototype structure. Remember that it has been established that some moral concepts such as CRIME, HIGHLY MORAL PERSON, RESPECT, and JUST PERSON have prototype structure. However, given that a prototype is a summary representation of features, it takes several experiences with instances of moral actions, cases, or encounters with virtuous agents to arrive upon a summary representation of, for example, JUST PERSON. Thus, there is a point in early childhood where we do not have enough experiences to form a summary representation. If we do not have a prototype for JUST PERSON yet in early childhood but we will at a later point, then at this earlier time, we must be relying on representations of specific agents or, in other words, conceptual *exemplars* for filling out the concept of JUST PERSON. In this fashion, we can see that there is a chain linking prototype theory to exemplar theory because if some moral concepts have prototype structure, then it is the case that at some earlier stage, moral

concepts such as those examined in the prototype theory section of this chapter, have exemplar structure.[27] Thus, prototype theory, by way of concept acquisition, has built into the theory the existence of exemplar structure at some earlier time.[28]

3.4 The Armstrong, Gleitman, and Gleitman Objection

For the remainder of this chapter, we will entertain two objections to the prototype and/or exemplar theories. A well-known objection against the prototype and exemplar views attacks the typicality effects that these theories are built to account for. Armstrong, Gleitman, and Gleitman have run experiments on mathematical concepts such as EVEN NUMBER that are thought to have classical structure.[29] However, they found that such concepts also showed typicality effects, where participants believed that numbers such as 4 and 8 are more typical of the class *even number* than 34 and 106. They argue that since there are typicality effects for concepts that people do have definitions for, such effects do not provide evidence concerning the structure of concepts. In other words, prototypes and exemplars are posited, in part, to account for typicality effects. Due to typicality effects, prototype and exemplar theorists, in part, have drawn

---

[27]This point that prototype theory acknowledges the viability of the exemplar view has been anticipated by the likes of Murphy and Wong. Gregory L. Murphy, *The Big Book of Concepts*, (Cambridge, MA: The MIT Press), 2002. Wong, Ibid.

[28]It is also noteworthy to mention that when one has had enough experiences with exemplars to form a summary representation, it is highly unlikely that one's exemplars of a moral category will simply disappear. While there may be a debatable issue in regards to our memory capacity for retaining exemplars, when a prototype is formed, it is also the case that we will still have exemplars as well.

[29]Armstrong, ibid.

structural conclusions regarding concepts. Yet, since EVEN NUMBER has classical structure and is composed of DIVISIBLE BY TWO WITHOUT REMAINDER rather than having prototype or exemplar structure and it also shows typicality effects, such effects are not an indication or pertinent to the determination of the structure of concepts. Therefore, typicality effects do not show that concepts have prototype or exemplar structure. This argument attacks one means for showing that the prototype and exemplar theory are viable views of concepts.

The problem with this objection is that Armstrong, et al. have not tested for whether EVEN NUMBER indeed does have classical structure or not. If eternal correct mathematical concepts rather than conceptions are really platonic abstract objects, it is the case that the platonic correct concept has a definition. Nevertheless, is it the case that the psychological conception of EVEN NUMBER in the minds of fallible human beings has classical structure? They have assumed without evidence that it does. However, given that EVEN NUMBER is an abstract conception, they have not provided evidence consistent with PAC that it does have such a structure, nor have they provided relevant data that it only has such a constitution. Therefore, we cannot confidently state that it does, and their objection is inconclusive barring further evidence.

For, it very well could be the case that there is a mathematical concept pluralism, where mathematical concepts have several viable structures open to them. It may be the case that there is no one grand theory of concepts in the mathematical domain and several concept theories are viable in this realm as well. For instance, Lakoff and Nunez, in

Where Mathematics Comes From, have already argued in a manner consistent with PAC that mathematical concepts are constituted by theory-theory metaphorical background knowledge.[30] For instance, Lakoff and Nunez claim that the metaphor, 'a set is a container,' when mentally represented, may play a role in the higher competences. Thus, they argue that SET is constituted by CONTAINERS, and NUMBERS are constituted by the metaphorical background knowledge relation of ARE OBJECT COLLECTIONS AND POINTS ON A LINE. Thus, for example, it may be the case that our concept of EVEN NUMBER may have classical, prototype, exemplar, and theory-theory structure. When given categorization tasks that test for typicality effects, our prototype or perhaps exemplar concept knowledge is used, but when given other tests, the classical concept knowledge may be at play. Much of this is mere speculation that is not based on evidence consistent with PAC, but since Armstrong, et al. have not shown that EVEN NUMBER has classical structure nor that it only has such a structure and there is evidence that mathematical concepts have a non-classical constitution, I am illustrating a possibility of how an overall pluralistic theory of mathematical concepts can account for Armstrong, et al.'s particular typicality findings. The bottom line is that in order for their objection to work, they need to provide evidence consistent with PAC that mathematical concepts such as EVEN NUMBER have definitional structure and that such concepts also do not have prototype or exemplar structure. So long as this has not been done, their objection cannot

---

[30] George Lakoff and Rafael Nunez, *Where Mathematics Comes From*, (New York: Basic Books), 2001.

come to fruition, and there is always the possibility that a mathematical pluralism may account for their findings and that typicality effects are indeed an indication of prototype or exemplar structure. Therefore, based on PAC, their contention against the prototype and exemplar theories and typicality effects is to date unfounded.

In response, they may claim that in their third series of experiments, they found that a different group of participants believed that EVEN NUMBER does have a strict definitional structure. They asked these new participants: "Does it make sense to rate items in this category for *degree of membership* in the category?"[31] The experimenters then defined what 'degree of membership' means:

> "It makes sense to rate items for degree of membership in a category if the items meet the criteria required for membership to a *different degree*. It does *not* make sense to rate items for degree of membership in a category if all the items meet the criteria required for membership to the *same degree*; that is, if the items are literally either in or out of the category."[32]

They found that most subjects believed that items that are fruits, vegetables, and vehicles can be rated by degree of membership, but all participants believed that members of the class *even number* cannot be rated by degree. This may suggest that all participants believed that their concept EVEN NUMBER has classical structure. Armstrong and company may argue that since subjects believe that EVEN NUMBER has a definition, it must be the case that it does in fact have a definitional structure. The problem with this is that this third study does not test for whether EVEN NUMBER has classical structure. For,

---

[31] Armstrong, 242.
[32] Armstrong, 242.

one can believe that one's concept has a particular structure, but it does not then follow that one's concept actually does have this structure. Empirical tests concretely may show that one's concept really does not have this structure but rather has some other constitution. For example, I can believe my concept of BELIEF has prototype structure, but Hampton's aforementioned test on abstract concepts in the first chapter indicates that I may be mistaken and that my BELIEF concept most likely may not have such a constitution. Thus, at most, all their third experiment shows is that participants have the belief that their concept EVEN NUMBER has definitional structure. This third experiment does not provide evidence that EVEN NUMBER actually has classical structure.

3.5 The Exemplar Theory Objection to Prototype Theory

For concrete concepts, several exemplar theorists have claimed the superiority of their own view over the prototype theory such that we may discount the latter view.[33] For example, the first of two problems, *inter alia*, for the prototype theory is that old familiar particular members are categorized more quickly and accurately than new potential members despite the fact that they both are equally typical members based on

[33]Medin and Schaffer, ibid. W. Wattenmaker, G. Dewey, T. Murphy, and D. Medin, "Linear separability and concept learning: context, relational properties, and concept naturalness," *Cognitive Psychology* 18: (1986), 158-194. G. Murphy and A. Kaplan, "Feature distribution and background knowledge in category learning," *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* 53A: (2000), 962-982.

the prototype view.[34] For instance, I can more easily classify my Rottweiler as a dog than a Dalmatian, Dalmatians being a kind of dog that I have never seen before, even though they both are equally typical. The prototype view claims that categorization is based on an item's similarity to a list of statistically frequent features rather than on previously experienced particular instances. Hence, if two items are equally typical members on the prototype view, then the old item should not be more easily categorized than the new item on this theory. However, since the old item is more easily classified as a member, then this provides support for the exemplar paradigm since ease of categorization based on this view depends on the similarity of an item to a particular stored instance of the category. Since the old item is directly represented by a stored exemplar but the new item is not, the exemplar theory can account for these categorization results.

A second problem is that a less typical member of a category as measured on the prototype theory can be more easily classified and can be learned more quickly as a token of the category as compared to a more typical member if the less typical member is more similar to previously experienced members of a class.[35] For instance, I can more easily classify a Chihuahua as a dog as compared to a St. Bernard because my sister owns a Chihuahua and I have never seen a St. Bernard before. This holds true despite the fact that the St. Bernard would be a more typical member of *dog* for me on the prototype view

---

[34] R. Nosofsky, "Exemplars, prototypes and similarity rules," in *From Learning Theory to connectionist Theory: Essays in Honor of W. K.* Estes, ed by A. Healy, S. Kosslyn, and R. Shiffrin, (Hillsdale, NJ: Erlbaum, 1992), p. 149-68.
[35] Medin and Schaffer, ibid.

than the Chihuahua. This phenomenon can be explained by the exemplar rather than prototype view since such results are influenced by the similarity of an item to a previously experienced particular instance of a class. Given that the exemplar view appears to outperform the prototype theory in numerous areas, some exemplar theorists then make the leap in categorically believing that our concrete concepts do not have prototype structure.

Thus, likewise for abstract moral concepts, it may be the case that exemplar theorists claim that moral concepts do not have prototype structure. Rather, despite some evidence for the prototype view for moral concepts, such mental representations really only have exemplar structure. However, the first problem with this potential objection is that although it may be possibly true, as of yet there is no evidence consistent with PAC to show that moral concepts have exemplar rather than prototype structure. While experiments in the concrete concept realm may show that such concepts have exemplar rather than prototype structure, there is no analogous specific study on moral concepts to demonstrate this conclusion. Such an experiment needs to be conducted in order to remain true to PAC and show that the structure of abstract moral concepts, like those of concrete concepts, is possibly constituted by exemplars rather than prototypes.

Second, there is recent evidence that prototype and exemplars both are used in cognition in different situations for concrete concepts. While all the numerous studies that attack the exemplar theorist's claim to dominance over the prototype theory made by

psychologists and philosophers cannot be reviewed here,[36] we will now examine several studies that indicate that both prototype and exemplar knowledge is used in cognition at different times depending on the circumstances. To note, it is very difficult to empirically distinguish between whether prototype or exemplar knowledge is being used since they both broadly speaking make similar predictions in regards to such things as there being typicality effects. Barsalou also has argued that the exemplar theory is more powerful than the prototype view since exemplars represent all the observed members of a class with their concomitant features, which is larger than a list of summary features.[37] Given that the exemplar view can mimic prototype theory since one's set of prototypes is contained within one's larger set of exemplars and the exemplars' components, it is difficult to distinguish which model is being used in cognition since even if subjects are using prototypes, an exemplar model can still account for the data.

---

[36] For further studies not presented in the text that show that both prototypes and exemplars are used in cognition, see: B.W.A. Whittlesea, L.R. Brooks, and C. Westcott, "After the learning is over: Factors controlling the selective application of general and particular knowledge," *Journal of Experimental Psychology: Learning, Memory, and* Cognition 20: (1994), 259-74. Weiskopf, ibid. D. Medin, M. Altom, and T. Murphy, "Given vs. induced category representations: Use of prototype and exemplar information in classification," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10: (1984), 333-52. T. Yamauchi and A.B. Markman, "Category learning by inference and classification," *Journal of Memory and* Language 39: (1998), 124-48. P. Juslin, S. Jones, H. Olsson, and A. Winman, "Cue abstraction and exemplar memory in categorization," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29: (2003), 924-41.

[37] Lawrence Barsalou, "On the indistinguishability of exemplar memory and abstraction in category representation," in *Advances in Social Cognition, Vol. III: Content and process Specificity in the Effects of Prior Experiences*, ed. by T. Srull and R. Wyer, Jr., (Hillsdale, NJ: Erlbaum, 1990), pp 61-88.

However, recall from the second chapter that the exemplar theory is better equipped for classifying atypical or odd member items than the prototype view given that exemplar theory concepts may actually be constituted by mentally represented atypical members of a class. On the other hand, since prototype theory concepts are only constituted by a summary representation of the features of members of a class, subject using such concepts will have a difficult time properly classifying odd member items; atypical items that generally do not match the summary features of the class. Common to a number of prototype versus exemplar experiments, it is with this fact in mind in which such experiments are designed in order to attempt to discern whether participants are using prototype or exemplar knowledge. In other words, one established way to distinguish between the uses of the two types of knowledge is by having atypical members in one's experiment that do not share many features represented by prototypes but, on the other hand, are highly similar to a particular token of the class. If subjects can classify the atypical items, then this provides evidence that the exemplar knowledge is being recruited.

One method used to distinguish between the prototype and exemplar theories is to use artificial categories. Artificial categories are usually a string of letters or numbers that are completely made up by experimenters, are meaningless, and are not intended to refer to any ordinary category used by human beings such as with natural and artifact kind concepts. For example, Rosch and Mervis used strings such as HPNWD and

JXPHM in a particular study to examine the prototype theory.[38] Artificial categories and their unusual properties also may allow some experimenters to more easily distinguish whether participants are relying on one structure rather than the other. Now, there are concerns of ecological validity for artificial categories, and studies using them may only be done with a cautionary note that such findings have not been tested yet for real categories.[39] While artificial category tests have been used for various purposes in relation to concepts, stronger inferences may naturally be made if such studies are replicated implicating the actual type of category in question.

Like several exemplar theorists, Smith, et al. have run artificial category studies pitting prototypes against exemplars.[40] However, they made two vital alterations to their design as compared to certain exemplar studies that illustrated the purported dominance of exemplars over prototypes.[41] First, Smith, et al. used artificial categories that had stronger structures than those used in the exemplar experiments. What is meant by stronger structure is that they used artificial categories with more features than the typical

---

[38]Eleanor Roschand Caroline Mervis, "Family resemblances: Studies in the Internal Structure of Categories," *Cognitive Psychology* 7: (1975), 573-605.

[39] For example, there are criticisms on the ecological validity of the numerous artificial category studies run by exemplar theorists. J.D. Smith, M.J. Murray, and J.P. Minda, "Straight Talk About Linear Separability," *Journal of Experimental Psychology* 23: (1997), 659-80. J.D. Smith and J.P. Minda, "Prototypes in the mist: the early epochs of category learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: (1998), 1411-36.

[40]Smith, Murray, and Minda, ibid.

[41] D. Medin and P. Schwanenflugel, "Linear separability in classification learning," *Journal of Experimental Psychology: Human Learning and Memory* 7: (1981), 241-53. W.D. Wattenmaker, G.I. Dewey, T.D. Murphy, and D.L. Medin, "Linear separability and concept learning: Context, relational properties, and concept naturalness," *Cognitive Psychology* 18: (1986), 158-94.

three to five features used in artificial category studies run by those such as Medin and Schwanenflugel. By having more features to an artificial category, such a study is more ecologically valid and is closer to ordinary categories in that ordinary categories such as *birds*, *dogs*, and *trees* can have a larger number of features. Second, instead of taking all the subject's data, grouping them together, and then determining whether prototype or exemplar theory is more dominant, Smith, et al. determined which concept theory is more dominant relative to each individual without pooling all the data together. In other words, they assessed whether prototype or exemplar structure was used on an individual-to-individual basis. What they found was that half the subjects utilized prototype retrieval in categorization while the other half used exemplar retrieval. In this study, prototype participants were so labeled because they were able to classify items that were similar to a summary representation. However, such subjects performed poorly in categorizing atypical items. On the other hand, exemplar participants performed well with atypical items. Later this study was replicated by Smith and Minda, with the additional finding that participants can shift their use of a concept structure to a different structure over periods of time.[42]

There is also experimental data using real categories that confirm the recruitment of both prototype and exemplars in different circumstances for categorization. For example, Machery and Weiskopf discuss Barbara Malt's study in which subjects were found to use both prototype and exemplar-based categorization judgments for groupings

---

[42]Smith and Minda, ibid.

of different animals.[43] For example, in the third of six experiments, participants were taught during a learning stage both the prototype and exemplars of categories by being presented with pictures of various animals. She found that a group *a* of participants had trouble learning the prototypical features of an animal category *x* during the learning phase. Meanwhile, a group *b* of participants had no trouble learning the prototypical features. After this learning phase, both groups were presented with categorization tasks testing them with new and different animal pictures that belong to *x*. She found that this categorization task significantly primed for previously learned exemplars of the category *x* in group *a* rather than group *b* subjects. This is so because after this categorization task, group *a* subjects could more quickly categorize the particular animal picture that belonged to *x* that was previously learned during the learning stage as compared to group *b* participants. This priming suggests that an exemplar categorization strategy was used by group *a* subjects when classifying new pictures. However, when certain subjects could easily learn the prototypical features of a category, the exemplar priming effect was much weaker. This suggests that in these cases, group *b* subjects were relying on prototypes when categorizing new animal pictures.

Therefore, given the host of studies showing the use of both prototype and exemplar knowledge in the concrete domain, the present objection against the viability of the prototype theory for moral concepts fails. Exemplar theorists may want to object that

---

[43] B.C. Malt, "An online investigation of prototype and exemplar strategies in classification," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15: (1989), 539-555.

since concrete concepts are made up of exemplars rather than prototypes, it is the case that moral concepts are so constituted as well. However, along with the fact that this argument violates PAC, the above studies show that it is not the case that concrete concepts are constituted by exemplars rather than prototypes. Rather, both theories are viable for concrete concepts.

4. The Theory-Theory and the Emotion Theory

In this chapter, I assess the viability of the theory-theory and the emotion theory for moral concepts. As the evidence for these views will rely on studies from moral psychology that show that theories and emotions play a causal role in moral cognition, I first analyze how one can use such causal studies in order to draw constitution claims on moral concepts. I contend that by relying on functionalism from the philosophy of mind, we may use the causal studies in order to draw constitution claims in certain cases. Then, I discuss when the use of causal studies in order to make constitution claims is appropriate and inappropriate. Upon finishing this analysis, I then put forth empirical causal evidence that the theory-theory is prima facie a viable theory of moral concepts. Next, I criticize Jesse Prinz's psychopaths-based argument that all moral concepts are constituted by sentiments and emotions. The objections to Prinz's contention open the pathway to show that using causal evidence rather than studies on psychopaths is the more successful way to argue for the emotion theory. After criticizing Prinz's argument, I contend for the prima facie viability of the emotion theory by drawing on causal studies from moral psychology.

4.1. Causal Moral Psychology Studies & the Constitution of Moral Concepts

As previously mentioned in the first chapter, concepts are functionally defined as mental representations that play a causal role in the higher cognitive competences of the mind, such as in categorization, induction, deduction, and analogical reasoning. In

arguing for the viability of the theory-theory and the emotion theory for moral concepts, such conclusions are motivated by the issue of categorization since given the evidence in moral psychology, at times theory-based knowledge and/or emotions play a specified and qualified kind of causal role in determining when *a* is categorized as an *F*.[1] In this and many other respects, empirical data can be used to show that theory knowledge and emotions can play an important qualified role in the relevant higher competences. However, Prinz distinguishes between causation and constitution in the causal moral judgment literature in empirical moral psychology, where philosophers and psychologists examine what factors influence moral judgment. Prinz warns that we cannot reach any constitution conclusions for moral concepts based on causal evidence. For instance, he writes, "The fact that emotions influence moral judgments does not entail that moral judgments contain emotions."[2] However, I interpret making moral judgments as acts of categorization. As previously illustrated in Machery's definition of concepts and as also pointed out by Weiskopf, in psychology concepts are thought to play a (specified kind of) functional or causal role in categorization and in other relevant higher competences.[3] Based on the functional definability of concepts, in certain cases they are constituted by those mental representations or structure(s) that best fills or realizes the specified kind of causal role. This is the underlying metaphysics in the concepts literature. Thus, in some

---

[1] The specified and qualified kind of causal role will be elaborated upon shortly.
[2] Jesse Prinz, *The Emotional Construction of Morals,* (Oxford: Oxford University Press, 2008), 28.
[3]Daniel Weiskopf, "The Plurality of Concepts," *Synthese* 169: (2009), 145-173. "Concept empiricism and the vehicles of thought," *Journal of Consciousness Studies* 14: (2007), 156–183.

cases if certain judgments are caused by theories, then the relevant concepts are constituted by theories since theories best realize the specified kind of causal role. Likewise, in some cases if certain judgments are caused by emotions, then the relevant concepts are constituted by emotions since emotions realize the specified kind of causal role.

Mental representations of the statistically frequent features of members of a class and mental representations of particular instances of a class causally explain how we may classify certain three-dimensional objects. Thus, the prototype and exemplar theory, respectively, may be spawned since prototypes and exemplars best fill the specified causal roles. Folk theoretical beliefs of biological hidden essences and causal laws in a specified way influences categorization of whether a biological object falls within a certain natural kind. Thus, the theory-theory of concepts may be a viable view and some biological natural kind concepts may be constituted by folk beliefs about hidden essences and causal laws. Theory knowledge can causally influence some classifications in the moral domain. What realizes the specified causal role in these relevant cases of these acts of classification are theories, and thus, the relevant concepts may be constituted by theories. Likewise, emotions can also influence moral judgments in an appropriate qualified way, so in such cases, the relevant moral concepts may be constituted by emotions.

This functional handling of concepts is analogous to what takes place with artifacts and many biological properties. An engine is functionally defined as playing a

causal role in transforming various forms of energy into mechanical force or motion, where an engine just *is* that which realizes the causal role. Thus, the thing under the hood of my car and in my lawnmower, in virtue of realizing the specified causal role of transforming energy into mechanical force, is an engine. A heart is functionally defined as an organ that plays a causal role in pumping blood. Thus, whatever realizes or plays the causal role such as my heart, a lizard's heart, or an artificial heart is a heart because it fulfills the causal role. Once we have functionally defined what a heart is and we have determined that what I call "my heart" realizes the causal role of pumping blood, then what I call "my heart" is indeed a heart. Likewise, as concepts are also functionally defined, with an eye towards explaining behavior, as that which plays a specified causal role in categorization *inter alia*, once we determine what best realizes the specified causal role (e.g., a prototype), a concept may be the relevant realizer. While many philosophers and scientists of the mind are well aware of the below causal studies, to the best of my knowledge, no one has been able to see the implications of such studies in light of the functionalist metaphysics of some moral conceptual mental states being theories and/or emotions.

However, not any causal influence in the higher competences can be said to constitute the relevant concept. For, this would allow too many different things to constitute one's concept; some mental representations that we ordinarily would not think are relevant conceptual components. For instance, an experimenter may give participants drugs before a natural kind categorization task that makes them feel the emotion of

sadness. This sadness may act as an influence at the front end of the categorization process temporarily biasing subjects' responses to be significantly different than if they had not been given the drug. Here, it seems that we would not say that subjects' LION concept is constituted by sadness even though sadness plays a causal role.

I will now introduce and state an important qualification concerning the fact that a causal influence can lead to a constitution claim on concepts. As I will explain, there are several characteristics that will allow us to determine when an influence is or is not constitutive of a concept, and a positive account of this will now be offered. We may claim that a causal influence generally leads to a concept constitution claim when the influence is knowledge that is used by default in moral reasoning and is at least in part inferable, deductively or inductively, from the concept in question's inferential base.[4] Hence, so long as a mentally represented influence is default knowledge that leads to or, in other words, consciously or subconsciously influences a sincere judgment in a certain moral situation *x* and it is at least in part inferable from the concept in question's inferential base, then this generally leads to a constitution claim on concepts. To note, I use the term 'generally' in the previous sentence because when it comes to emotions being in part inferable from the inferential base, an alteration to the above rule of when

[4] This is assuming that the influence is not itself the inferential base. Also, recall that by 'knowledge' I mean an information carrying mental state. The knowledge one has stored in a concept may be false, and emotions can also be considered to be knowledge so long as they are representational.

an influence actually constitutes a concept must be made. This alteration will be discussed later in this section.

Remember from the first chapter that concepts are thought to contain default bodies of knowledge rather than all the possible facts we may know about *x*. Default knowledge is knowledge that preferentially and presumptively is used in most of the higher competences. This allows for a significant reduction in the systematic selection of relevant knowledge when we generally reason about *x*. Furthermore, as countless thoughts and beliefs potentially may be in part inferable from the inferential base, having only default knowledge constitute a concept cuts down on the many different pieces of knowledge that may potentially constitute a moral concept to only those that play a main role in moral deliberation and in how we reach moral conclusions.

An inferential base is a theory-theory body of knowledge from which other mentally represented influences may be thought to be explained.[5] As illustrated in the second chapter, it is theory knowledge of normative ethical theories that explains and systematizes other more specific bodies of knowledge. Insofar as ethical theory knowledge provides an explanation for why we hold, for example, the prototype

---

[5] First, recall from the third chapter that the classical view is not viable for moral concepts. Thus, the inferential base is not a classical body of knowledge. Second, we may think of influences that actually do constitute the concept as being criterial in the simple sense that what constitutes a concept is what the mind actually uses consciously or subconsciously in a qualified way in categorization, reasoning, etc. A concept is functionally defined as being constituted by that which plays the specified role in categorization, reasoning, etc. In this case, influences need not consciously be mentally represented as being criterial. Things may remain at the subconscious level.

knowledge that we do, such ethical theory knowledge falls within the confines of the theory-theory. Furthermore, the inferential base is the most firmly held theory belief(s), such that it would be the last theory belief to go in case one's system of knowledge required revision. The inferential base is specifically a theory belief(s) that occupies a defeasible placeholder position, where this theory belief(s) is specifically the most firmly held out of the set of theory beliefs one may have.[6] The inferential base contains the most firmly held theory belief because we want to be sure that the causal influence can lead to a constitution claim for the relevant concept. If whether an influence constitutes a concept generally and in part turns on whether it may be in part inferable from theory knowledge since theory knowledge systematizes and explains other kinds of knowledge, it generally should be in part inferable from the most firmly held and stable theory belief in order to be certain that the influence may be said to constitute the concept. This is why the inferential base only contains the most firmly held theory belief. The theoretical work that the inferential base performs is that if a default influence is in part inferable from the base, then the influence constitutes the relevant concept.

Now, I make the qualification that the influence generally must be at least in part inferable rather than just merely inferable from the inferential base because there may be many subsidiary assumptions needed along with the inferential base in order to draw the

---

[6] It may be the case that a theory belief may be the first to go under certain pressures and epistemic contexts, but it is the last to go under different pressures and contexts. If this is the case, then the inferential base is theory knowledge that holds up in more contexts than other theory knowledge. If there is an even split between two theories where they hold up in relation to different but an equal number of contexts, then both theories are part of the inferential base.

relevant inferences. If one claims that an influence constitutes the concept when it generally is inferable from the inferential base, then this means that the inferential base must also contain all of these subsidiary assumptions, which may lead to the inferential base having to contain an overwhelming amount of one's knowledge. We obviously do not want the inferential base to contain all of this knowledge, so we must find a way to delimit what knowledge is in the inferential base. This specifically is a variation of the holism problem for the theory-theory in that the theory knowledge that constitutes one's relevant concept unreasonably may balloon out to inevitably include all of one's total knowledge. The claim that an influence must be inferable from the inferential base will run into the problem of holism. However, we may escape this holism problem by claiming that the inferential base contains only theory knowledge that is the last to go in case one's system of knowledge requires revision. It does not contain all of one's knowledge. Moreover, an influence need not be inferable from only what is contained in the inferential base in order to constitute the relevant moral concept, but rather, an influence generally need only be *in part* inferable from the inferential base, where subsidiary assumptions also may play a role in the relevant inference, but such ancillary assumptions need not themselves constitute the inferential base. This is due to the qualification that one criterion for constituting a concept is that an influence generally need be only in part inferable from the inferential base, an inferential base that only contains the most firmly held theory knowledge, but other ancillary knowledge that does not lie in the base may also play a role in the inference. Only when an influence is

default knowledge and generally is in part inferable from the inferential base, a base that only contains the most firmly held theory knowledge, can we then say that the influence constitutes the concept.

As an example of an inferential base, let us assume there is an agent named Christian whose inferential base belief is OBEY THE COMMANDS OF GOD. This is theory-theory knowledge of divine command theory, where his moral prototype and exemplar knowledge, such as DO NOT STEAL, JESUS HELPING MARY MAGDALENE, and DO NOT LIE, may be thought to be in part inferred from this inferential base. To be sure, there are subsidiary assumptions needed along with the inferential base in order to reach the inferences that, for instance, one should not steal and lie. For example, one may need premises like 'Moses led his people out of Egypt and came upon a burning bush,' 'the burning bush was God,' 'God spoke to Moses through divine revelation,' and 'God gave Moses the Ten Commandments.' However, such subsidiary premises need not be represented in Christian's inferential base for his theory-theory MORAL concept since inferential base knowledge only contains the most firmly held ethical theory knowledge and generally, in order to constitute the relevant moral concept, one criterion is that inferences such as DO NOT STEAL need not be inferred solely from what is contained in the inferential base. Rather, in order to constitute the concept, an influence used by default need only be at least in part inferable from the inferential base. In other words, an influence generally needs to in part be inferred from the inferential base, but ancillary premises not contained in the inferential base can also play a role in the inference. Being

in part inferred does not mean that an influence must be solely inferred from the inferential base. Here, the above subsidiary assumptions regarding the story of Moses can be mentally represented constituents of Christian's BOOK OF EXODUS or STORY OF MOSES AND GOD compound concepts rather than being a part of his firmly held theory knowledge contained in his MORAL concept. This is how one may handle the holism problem for the inferential base.

As it will be demonstrated later in this chapter that emotions can also at times constitute our moral concepts, it may initially be thought that emotions may be at least in part inferable from the inferential base when emotions work jointly with or perhaps in some cases are in part constituted by relevant cognitive bodies of knowledge; normative cognitive knowledge that helps to specify that a judgment that is in part constituted by an emotion is a moral judgment rather than a non-moral judgment.[7] As examples of how certain emotions may be thought to be in part inferred from the inferential base, feeling joy when given the opportunity to help others and feeling guilt when one has done something unethical may be thought to all be in part inferable from the inferential base OBEY THE COMMANDS OF GOD. Moreover, this inferential base would be the last belief to go when Christian's system of knowledge requires revision. For example, by hypothesis, since his inferential base is constituted by OBEY THE COMMANDS OF GOD,

[7] Recall from the second chapter that I leave open the possibility that some but not all emotions in part may be constituted by cognitive constituents.

when questioned by a skeptic to revise his beliefs, the last belief Christian would let go of is OBEY THE COMMANDS OF GOD.

However, strictly speaking, emotions that play a causal role, completely as such, are not in part inferable from the inferential base. Rather, in our above examples, something like mentally represented prototype rules that merely talk about emotions rather than actually being an experience of an emotion, such as ONE POTENTIAL FACTOR FOR SEEING SOMETHING AS A WRONG ACT WOULD BE IF THE PERFORMING OF THE ACT ELICITS GUILT, is in part inferable from the inferential base since the Bible at least implicitly says that the feeling of guilt in many cases influences people to make judgments that their act in a situation is wrong. The emotion, completely as such, is not in part inferable from the base. For example, if the emotion of guilt contains a cognitive component such as HARMING OTHERS DECREASES THEIR WELL-BEING along with other components such as qualia, the cognitive component may be in part inferable from the base, but the quale is not. The raw feel of a quale cannot be deductively or inductively in part inferred from the base. Rather, strictly speaking, only cognitive aspects of the emotion can be so inferred from a cognitive base since the base only contains cognitive mental states from which inferences may in part be drawn. Therefore, the emotion, completely as such, is not in part inferable from the base. If the emotion of joy is non-cognitive, then strictly speaking, it also is not in part inferable from the base since it is non-cognitive.

In order to resolve this issue of when an emotional influence is a correct or appropriate influence that actually constitutes the relevant concept, we may state that only for emotional influences, an emotional influence constitutes a concept when 1) the emotion is used by default and 2a) for an emotion with a cognitive constituent, at least one of the cognitive constituents is in part inferable from the inferential base. For instance, since Christian's cognitive emotion of guilt is used by default when categorizing things under WRONG ACT and this emotion's constituent HARMING OTHERS DECREASES THEIR WELL-BEING is in part inferable from Christian's inferential base OBEY THE COMMANDS OF GOD, guilt actually constitutes Christian's concept WRONG ACT. 2b) For a cognitive or non-cognitive emotion, a corresponding cognitive prototype rule must be held consciously or subconsciously by the agent, such as ONE POTENTIAL FACTOR FOR SEEING SOMETHING AS A WRONG ACT WOULD BE IF THE PERFORMING OF THE ACT ELICITS GUILT, that states that the expression of the influencing emotion in the type of moral circumstance at hand is a potential factor that may lead one to make the relevant categorization. Furthermore, this rule must be able to be in part inferable from the inferential base. Notice that this cognitive rule does not itself contain the feeling of the emotion, but it merely talks about the emotion. For instance, the feeling of guilt constitutes Christian's WRONG ACT concept in the relevant scenario since it is used by default when he categorizes acts as being wrong and his corresponding rule ONE POTENTIAL FACTOR FOR SEEING SOMETHING AS A WRONG ACT WOULD BE IF THE PERFORMING OF THE ACT ELICITS GUILT is in part inferable from his inferential base. The influence of guilt appears to be an appropriate or correct influence for constituting

the concept in this particular circumstance for Christian in significant part due to the fact that his corresponding cognitive rule about guilt is in part inferable from the inferential base. This is somewhat similar but not exactly similar to how one significant criterion for prototypes and exemplars is that they must be in part inferable from the inferential base in order to be proper influences that really do constitute the relevant concept.

Thus, the actual non-cognitive emotional influence of guilt in part constitutes Christian's concept WRONG ACT because for Christian, the emotion itself is used by default when categorizing things as wrong acts and his corresponding cognitive rule ONE POTENTIAL FACTOR FOR SEEING SOMETHING AS A WRONG ACT WOULD BE IF THE PERFORMING OF THE ACT ELICITS GUILT that talks about the emotion of guilt and that states that the expression of the influencing emotion in the type of moral circumstance at hand is a potential factor that may lead one to make the relevant categorization is in part inferable from Christian's inferential base OBEY THE COMMANDS OF GOD. Therefore, the actual emotion of guilt constitutes Christian's WRONG ACT concept. To note, cognitive emotions need to only satisfy either 2a) or 2b), along with 1), in order to be a proper influence. It does not have to satisfy both 2a) and 2b). However, it is perfectly possible that there are cases where a cognitive emotion may indeed satisfy both 2a) and 2b).

In what follows, purely for ease of linguistic expression, when referencing the overall general conditions for when an influence actually constitutes a concept, I will state the above original conditions for when non-emotion influences constitute a concept. However, this is not meant to in any way abnegate the special conditions just discussed

particularly for emotions. I will discuss and rely on the alternate emotion-based conditions when *only* emotional influences are being discussed.

Now that we have stated the characterizations of when an influence does constitute the relevant concept, we now will discuss several tests that may be used to help us make stronger inferences as to whether an influence actually constitutes the concept. These additional tests may be run after one has already conducted initial studies determining that a particular influence(s) may constitute the concept in question. For example, these tests may be conducted after a theory-theory study is run in which a particular ethical theory knowledge that systematically explains one's prototype and exemplar knowledge is initially shown to influence a categorization judgment. One such additional test is seeing whether an agent does not retract a judgment that was issued from a particular influence. If the judgment is not retracted, then this suggests that the relevant influence may constitute the concept. Call this the *retraction test*. Here, it is vital that the retraction test be conducted in normal conditions when the experimenters are not attempting to in some way manipulate the participant by trying to get the agent to make a mental error during the test. For instance, the participants should not be hypnotized or be induced with mood altering drugs. Based on the retraction test, if the influence is later retracted, then it is suggestive that the influence does not constitute the relevant concept. It is important to emphasize that the retraction test is not meant to be a decisive method for determining whether a particular influence is in fact the constituent of a concept or not. It is only a test that provides merely suggestive evidence that an

influence may be a constituent of a concept. Notice that it may not necessarily be the case that if the retraction test is passed, then the influence must be a constituent of the relevant concept. For, it may be the case that an individual is frequently prone to some kind of mental error and never retracts a judgment or mostly does not retract a judgment that issues from an influence that does not really constitute the relevant concept. One may wonder how to tell the difference between whether an unretracted judgment or a mostly unretracted judgment is a reflection of mental error or is a reflection of being influenced by a mental representation that is legitimately a constituent of the relevant concept. An unretracted or mostly unretracted judgment may be a result of mental error if there is a plausible causal explanation for why the judgment may be an error, such as when a doctor states that a patient is constantly making delusional judgments because the patient is under the influence of medication.

It is also important to make sure whether or not the influence in question or even judgment itself is actually part of the inferential base. For, if the influence is a part of the inferential base, then we may understand the ensuing judgment to not be a result of mental error because such a judgment was influenced by the inferential base knowledge. Also, if in the appropriate moral context, a judgment itself is merely a statement of one's most firmly held theory belief, then it also is not a result of mental error. One may determine whether a mental representation is a part of the inferential base or not by uncovering what someone's inferential base is. One may run theory-theory studies, which will be discussed in the next section, to see what ethical theory knowledge an

individual possesses and then also see what theory knowledge will be the last to go in case the individual's system of beliefs requires revision. If an individual maintains more than one ethical theory, then objections can be presented against these various theories and against the rules, reasons, virtues, etc. that may be in part inferred from these theories in order to see which ethical theory knowledge will be most firmly held when the subject's set of beliefs may require revision. Upon finding out what the inferential base is, one may then see whether or not the influence or judgment is a part of the base. All in all, even though the passing of the retraction test may not prove to be definitive, the passing of this test may be said to still provide a degree of suggestive evidence for an influence constituting a concept.

Some factors that may be at play for why an influence is still used in cognition even though, for example, it is not in part inferable from the concept in question's inferential base are various kinds of mental error or various causal routes to mental error such as: being hit on the head, mistaken reasoning, mental fatigue, being under the influence of drugs, etc. This appears to be a grab bag of factors, but they are all grounded in the fact that they are various forms of mental error or causal routes to such error. Moreover, such factors may potentially lead one to use a particular mentally represented influence that is really not a part of one's relevant concept.

As an example of a retraction, Christian may be momentarily influenced by the thought PROTECT MY LIFE AT ALL COSTS when deciding to retaliate against a random murderer who has intruded his home. However, Christian is Amish and believes in a life

of non-violence. Given Christian's fundamentally held theory beliefs such as OBEY THE COMMANDS OF GOD, we can understand why Christian later retracts his judgment to retaliate against the murderer because his inferential base or fundamental beliefs from which some of his other beliefs may be in part inferred and explained indicate that PROTECT MY LIFE AT ALL COSTS is not supported by them. This provides a deeper explanation for why retractions occur. For instance, Christian's later thought ALWAYS ABSTAIN FROM VIOLENCE influences his judgment to actually not physically defend himself against the intruder. Since this principle is at least in part inferentially supported by OBEY THE COMMANDS OF GOD and by hypothesis, it is default knowledge, this means that it in part constitutes Christian's concept MORALLY OUGHT.

As a qualification, when dealing with the retraction or unretraction of sincere judgments for constitution claims for only prototype, exemplar, and emotion theory knowledge,[8] we must state that the test for retraction is based on what constitutes Christian's relevant moral concept(s) as the inferential base at the time period the judgment is made t1, where there is no change or additions to his inferential base for his relevant moral concept(s) during time periods. For it is always possible that at a later time period t2, Christian may receive some kind of new moral education or go through some kind of radical moral conversion that changes what constitutes his relevant inferential base, even though the inferential base knowledge was previously the most

[8] The theory-theory is excluded because if inserted, it would lead to a circular statement: we apply the retraction test to determine what is part of the inferential base, but here we appeal to what is in the inferential base to decide how to apply the test.

firmly held theory knowledge. For example, at t2, Christian is influenced by an atheist book he reads and becomes an atheist. This effectively changes his inferential base for his concept MORALLY OUGHT. At t2, he has the principle USE FORCE TO DEFEND ONE'S OWN LIFE which leads him to retract the relevant previous judgments made in t1 in which he decides not to protect himself against violent intruders. Due to the event that kicks off t2, many previous moral judgments would be retracted, but it seems odd to say that during t1, Christian's concept MORALLY OUGHT is not constituted by ALWAYS ABSTAIN FROM VIOLENCE. For, this influence leads to continuous unretracted judgments until his moral conversion decades later such that there is suggestive evidence that it was in part inferable from the inferential base and it was used by default. So long as Christian's concept MORALLY OUGHT has the same inferential base constitution during a time period, his moral concept, generally speaking, is constituted by influences that are supported, explained, and systematized at least in part by the inferential base and that are used by default within that period of time but not outside that time period. Once there are additions or alterations to the inferential base of Christian's concept MORALLY OUGHT at the later period of time t2, this does not impugn the concept constitution claims that were made during t1.

Furthermore, a second possible test for figuring out whether a stronger constitution claim may be made in these types of cases is to see what a subject's inferential base is. Recall that I have already described above one possible way for determining what one's inferential base is. In finding what the fundamental and most

firmly held beliefs are, we may examine generally whether or not a given influence is in part inferable from them and whether it potentially constitutes the concept. If it generally is in part inferable from the inferential base, then this provides some suggestive evidence that the influence in question may constitute the relevant moral concept. Call this the *inferential base test.* This second test may show whether an influence in question, generally speaking, is at least in part inferable from an inferential base by first discovering what the inferential base is, but notice that it does not necessarily demonstrate that an influence is used by default. The inferential base test is not meant to provide definitive evidence that an influence in fact constitutes a moral concept. Rather, it is merely a suggestive test for concept constitution. For, it is possible that this test can show, in general terms, that an influence in question is at least in part inferable from the inferential base, but this influence still may not be used by default. Although not definitive, the inferential base test still lends some evidential support for drawing a stronger conclusion for when an influence actually constitutes a concept.

Knowledge of an agent's inferential base can also at times eliminate contradictory judgments. Let us return to Christian being attacked by the murderer before his conversion to atheism. Christian may at one point judge both that he should defend himself against the murderer and it is not the case that he should do so. However, the fact that one of his fundamental and most strongly held thoughts is OBEY THE COMMANDS OF GOD, this provides reason to believe that he does not really hold contradictory judgments in this particular case because the influence PROTECT MY LIFE AT ALL COSTS most likely is

not used by default given that PROTECT MY LIFE AT ALL COSTS is not at least in part inferable from the inferential base. Rather, he really believes that he should not defend himself against his assailant. Nevertheless, as one example of how an agent may really maintain contradictory beliefs, sometimes it may be possible that an agent maintains two opposing beliefs in his inferential base that are held with equally strong conviction. For instance, Christian* sincerely holds OBEY THE COMMANDS OF GOD as well as a kind of virtue ethical theory, in which one must exemplify the virtues of virtuous agents or perform those actions virtuous agents would perform, where one of the primary virtuous agents for Christian* is Malcolm X. Both theories for Christian* are equally important fundamentally basic beliefs. Due to this fact, Christian* judges that he should defend himself against the murderer, and he also judges that it is not the case that he should do so. Beliefs that are held with the same strong conviction within the inferential base leaves open the possibility for contradictions in that what directly influences his contradictory judgment is the equally weighted principles ALWAYS ABSTAIN FROM VIOLENCE and USE FORCE TO DEFEND ONE'S OWN LIFE. Here, we may say that Christian's concept MORALLY OUGHT is constituted by both of these equally weighted mentally represented principles since they are in part inferable by the equally firmly held inferential base beliefs and by hypothesis, they are used by default.

Moreover, a third test is to see whether the ensuing judgment from an influence still holds after the influence or the influence's informational value is made aware to the agent. If the judgment still holds, then this may also provide additional support for the

fact that the relevant influence constitutes the concept in question. Call this the *informational value test*. For example, in a study with eighty-four participants, most test subjects will give lower life satisfaction ratings when there is rainy weather, reflecting the impact of the weather on their current moods.[9] However, the negative effect of mood on life satisfaction ratings is generally eliminated when experimenters inquire about the weather to participants, which draws respondents' attention to this extraneous source of their mood. In these kinds of cases, the informational value test is a further suggestive way to see whether the influence actually constitutes the concept, although it may not necessarily show that this is the case. For example, a participant in the above study may still be influenced by the weather even when made aware of the weather by the experimenter. In a completely different example in which the informational value test also does not work, one may have a subconscious influence that does not really constitute the concept at hand, but this subconscious influence still persists and plays the relevant functional role even when the subject is made aware of the informational value of this faulty subconscious influence. As we can see, the informational value test is a mere test that may provide stronger suggestive evidence as to whether an influence may constitute the relevant concept. It is not meant to be a definitive means for determining whether an influence in fact constitutes the relevant moral concept. The participants' reflective access that is involved in the informational value test is not meant to be a definitive way

---

[9] N. Schwarz and G. Clore, "Mood, misattribution, and judgments of well-being: informative and directive functions of affective states," *Journal of Personality and Social Psychology*, 45: (1983), 513-523.

of seeing whether a given mentally represented influence in fact constitutes a concept. Rather, as can be seen in some instances of the above life satisfaction/rainy weather experiment, it provides defeasible suggestive evidence regarding concept constitution.

Hence, we can see that there are several suggestive ways to test for and make a stronger claim that an influence constitutes a concept, where an influence generally constitutes a concept based on the fact that it is default knowledge and it is at least in part inferable from the inferential base. To note, as merely a descriptive developmental point, influences that do constitute one's concept are usually arrived upon through intense social learning, childhood development, moral education, cultural influence, and interaction with one's family, friends, and peers. Through such external factors, others can influence, correct, and shape our moral concepts, and within a community, due to such interaction, it is the case that many members to a certain extent share some components of their moral concepts. It is important to note that social and environmental factors can play an important role in developing one's moral concepts and in determining what actually constitutes one's moral concepts.

Furthermore, so long as a participant in a study is not initially thought to be experiencing a possible state – such as being hypnotized, being hit on the head, being in stormy weather, or being under the influence of drugs – that may lead to mental error and the use of an influence that is not really a part of the relevant concept, a successful causal influence experiment has the strength to allow for an initial or prima facie positive claim that the influence in question does constitute the relevant concept. It will be understood

that such successful studies will allow for prima facie concept constitution claims. This may be the case even though the three retraction, inferential base, and informational value tests have not been conducted on the influence and corresponding judgment in question. However, the actual running of these three tests provides the additional support for making a stronger claim that the influence in question makes up the relevant concept. While the passing of the inferential base test provides the most certainty, the three tests along with the other possible various combinations of them can provide more support for making a stronger constitution claim as compared to making a prima facie claim. The strength of a claim is determined on a case by case basis taking into account the nature and design of the experiment in question as well as taking into account which of the three tests are and are not run. To note, most concept constitution conclusions drawn from particular experiments in the concrete concepts domain are prima facie claims.

4.2 The Arguments for the Theory-Theory

Now that we have clarified when we may rely on appropriate causal studies to draw constitution claims on moral concepts, we eventually will rely on causal evidence in order to demonstrate that the theory-theory is prima facie a viable theory of moral concepts. Goodwin and Darley ran experiments asking subjects to make judgments on a number of moral cases.[10] For example, subjects were asked whether they agreed or

---

[10] Geoffrey Goodwin and John Darley, "The psychology of meta-ethics: Exploring objectivism," *Cognition* (2007).

disagreed with the statement: Before the 3rd month of pregnancy, abortion for any reason (of the mother's) is morally permissible. Subsequently, participants were asked to identify the reason or justification for why they held the moral beliefs that they do. Close to a third of the subjects chose a supreme being or divine command theory as ordaining and grounding their own moral systems. Given that divine command theory is an ethical theory, this provides at least correlational evidence that there is a link between moral judgment and divine command theory for some people. While there is the possibility that such a potential theory-theory justification may be a post-hoc rationalization, where further studies need to be conducted in order to ascertain whether actual causation in the appropriate direction is involved, such an experiment does establish correlation.[11] Such correlational evidence needs to be coupled with relevant causal studies in order to make a sufficient claim for the viability of the theory-theory.

Kohlberg's stage theory of moral development also provides correlational support for a kind of folk consequentialism ethical theory, amongst others, that many individuals

---

[11] Jonathan Haidt has run a study asking participants if it is morally permissible for a brother and sister to secretly have protected sex on a one-time basis. While some participants stated that it was morally permissible, most claimed that it was morally wrong. When asked for the justification for their negative response, some said that it would lead to a deformed child. However, experimenters then reminded them that there was no chance to have a child given that they took the necessary precautionary measures. Some said that it may cause problems in their family, but experimenters reminded them that the sexual encounter was kept as a secret and the siblings even grew closer together as a result of the act even though they never had sex again. Haidt found that most of those who said the act was morally impermissible still held on to their judgment even though all the false non-moral facts they held in regard to the case were corrected. This study as well as others indicate that at many times reasons as well as non-moral factual justifications for belief, whether they may be true or false, may be post-hoc. Jonathan Haidt, "The Emotional Dog and its Rational Tail," *Psychological Review* 108: (2001), pp. 814-34.

may mentally represent. It provides only correlational support because Kohlberg gathered his evidence by interviewing his participants and asking them what justified their moral judgments when given hypothetical scenarios. Such purported justifications only may be post hoc rationalizations rather than being what causes participants to make judgments. Now, as it has already been argued in the third chapter that the classical view of concepts is not viable due to pervasive typicality effects for moral concepts as well as due to other reasons, we may, as has been pointed out in the second chapter, infer that if folk consequentialism knowledge does indeed constitute moral concepts, in this circumstance they are theory rather than definitional bodies of knowledge that occupy defeasible placeholder positions.

For Kohlberg, the first and second stage of pre-conventional morality is where children are egoistically concerned and influenced with rules in relation to punishment and self-interest. This supposedly provides evidence for a correlational link between moral judgments and a kind of ethical egoism theory knowledge. The third stage is where individuals are concerned about other people and their feelings. Moreover, agents are motivated to follow rules and expectations. Stage four of development is the folk consequentialism stage and is described as where the right is upholding the social order and maintaining the overall welfare of society or the group.[12] Here what duties and rules one must obey are determined by what the best overall consequences are for the group and by what upholds the social order, where each member's welfare is counted equally.

---

[12] Kohlberg, ibid, 410.

Each group member cannot egoistically do as they please. For, this may lead to chaos and disorder. The last two stages are post-conventional in that for individuals it is right to uphold the basic tenets and obligations stemming from the idea of a social contract and finally, in the last stage, to be guided by universal ethical principles that everyone should follow. Now, Kohlberg's stage theory has been heavily criticized, where, for examples, studies by the likes of Shweder and colleagues have shown that folk consequentialism is the most prevalent ethical theory knowledge that adults pervasively supposedly use across cultures.[13] Here, we may utilize these critical studies to conclude that there is a strong correlation between moral judgment and such a folk consequentialist theory knowledge.

We may understand the possible use of this consequentialism ethical theory as being one of *folk* consequentialism in that philosophers who are consequentialists tend to view this theory as a criterion of rightness rather than being a decision procedure that actually must be used in everyday circumstances. In other words, a consequentialist calculation is for determining which acts actually are right or wrong, but it should not be used to make everyday decision in all cases. For example, consequentialists generally believe that one should visit one's mother in the hospital not for the reason that it leads to the best overall consequences but because one love's one's mother. A consequentialist calculation is a criterion for determining that visiting one's mother is ethically right, but

---

[13] R. Shweder, M. Mahapatra, and J. Miller, "Culture and Moral Development," in *The Emergence of Morality in Young Children*, ed. by J. Kagan and S. Lamb, (Chicago: University of Chicago Press), 1987.

in real life, one must justify such an action by using alternative reasons. However, *folk* consequentialism will be understood as where the folk actually do use a consequentialist calculation as a decision procedure.

There are other experiments that actually may support the underlying causal influence of folk consequentialism on individuals' moral judgments, although this support is of a somewhat tentative nature. Joshua Greene, et al. ran a recent cognitive load study where subjects were filling out moral questionnaires on a computer.[14] They were presented with "high conflict" moral dilemmas in which subjects were asked whether it is appropriate to harm another individual in order to save several lives. One example of a high conflict dilemma that was used is the crying baby case:

> Enemy soldiers have taken over your village. They have orders to kill all remaining civilians. You and some of your townspeople have sought refuge in the cellar of a large house. Outside you hear the voices of soldiers who have come to search the house for valuables.
>
> Your baby begins to cry loudly. You cover his mouth to block the sound. If you remove your hand from his mouth his crying will summon the attention of the soldiers who will kill you, your child, and the others hiding out in the cellar. To save yourself and the others you must smother your child to death.
>
> Is it appropriate for you to smother your child in order to save yourself and the other townspeople?[15]

While answering such questions on moral dilemmas, numbers continually stream across the bottom of the screen and participants have to press a button when they see the

---

[14] J. Greene, S. Morelli, K. Lowenber, L. Nystrom, J. Cohen, "Cognitive load selectively interferes with utilitarian moral judgment," *Cognition* 107: (2008), 1144-54.
[15] Ibid.

number five. The result is that subjects selectively had a longer reaction time when making apparent folk consequentialism judgments – judgments that are based on taking into account the overall consequences for the group rather than only the welfare or rights of a single individual – under the cognitive load as opposed to having no cognitive load, but there was no increase in reaction time for supposed non-consequentialism-like judgments under cognitive load. This study provides causal rather than correlational support that purported folk consequentialism knowledge influences some moral judgments for some people because the longer reaction time suggests that the cognitive load of having to press a button when seeing the number five interferes with a controlled cognitive process, such as some kind of cost-benefit analysis reasoning process, whereas the cognitive load should have no effect on a fast automatic process. This conclusion is further buttressed by neuroimaging studies that show a strong correlation between the making of supposed folk consequentialism judgments and activation in the dorsolateral prefrontal cortex and inferior parietal lobe.[16] These brain regions are known for such things as complex planning, deductive/inductive reasoning, and long-term economic decision-making. In addition, in the next section on the emotion theory, we will examine several replicated trolley experiments that demonstrate that supposed folk consequentialism reasoning influences many participants' judgments.

---

[16] J. Greene, L.E. Nystrom, A.D. Engell, J.M. Darley, and J.D. Cohen, "The Neural Bases of Cognitive Conflict and Control in Moral Judgment," *Neuron* 44: (2004), 387-400.

Now, it may not be clear that Greene's study as well as the others definitively show that some individuals have supposed folk consequentialism knowledge. For, it very well could be the case that individual's might be relying on some other principle such as: *Bring about the best overall consequences so long as there is no self-interested cost to oneself.* This differs from folk consequentialism because the consideration of consequences is restricted by a self-interest clause, whereas folk consequentialism is not so restricted. Notice that this alternative principle also requires some calculation and is consistent with the greater reaction time that is indicative of a cognitive cost-benefit analysis process. Therefore, taken alone, it cannot be definitively stated that some people's moral concepts are prima facie constituted by folk consequentialism knowledge.

Although there is somewhat unclear causal evidence for a kind of folk consequentialism in this cognitive load study as well as in several other replicated studies to be discussed in the following section, these studies coupled with the cross-cultural correlational study on Kohlberg's folk consequentialism and the aforementioned neuroimaging studies allows us to prima facie infer that a kind of Kohlbergian folk consequentialist theory generally influences and at least in part constitutes adult moral judgments at times. Recall that a Kohlbergian folk consequentialist theory is where one must choose those actions that best maximize the overall welfare of the group or society, and the best overall welfare for a group is cashed out in terms of what best maintains societal or group order and functioning, where each member's welfare is counted equally. Moreover, on this view, individuals cannot do as they please or do what is in their best

interest in spite of the group or social order because this may lead to chaos and the potential destruction of the group. Since there is causal evidence in Greene's study as well as in others that some kind of apparent cost/benefit calculation at times influences moral judgment, this demonstrates that this kind of unspecified overall cost/benefit calculation influences judgment and is not a post hoc rationalization. Given that this is now on the table, recall that there is a correlation between the making of such judgments with activation in parts of the brain associated with planning and long-term economic decision-making, and there is a strong and unambiguous correlation between adult moral decision-making in a number of instances and a Kohlbergian folk consequentialism. With the correlational evidence in hand, we can now draw a conclusion that since some kind of cost/benefit analysis at times prima facie causally influences moral decision-making in a qualified way, this kind of cost/benefit analysis in many instances may be a Kohlbergian folk consequentialism knowledge. Moreover, as this kind of knowledge appears to be at work in a variety of moral situations as instanced by the given psychology experiments, it appears that this knowledge is used by default. Thus, there is a prima facie case that Kohlbergian folk consequentialism knowledge constitutes some moral concepts. Combining all the evidence allows us to make a novel prima facie claim that the theory-theory is a viable view for moral concepts and that some moral concepts for many people are constituted by a Kohlbergian folk consequentialist theory knowledge.

Regarding concept acquisition, recall from the third chapter that the concept acquisition and categorization desiderata can be thought of as two sides of the same psychological phenomenon – our disposition to put individual items into classes. If we have acquired a concept, then we can categorize items as members of the extension of the concept. Likewise, if we can use a concept to categorize, then at some point the concept has been phylogentically or ontogenetically acquired. Hence, we may infer that we generally acquire folk consequentialist theory concepts due to the above evidence for folk consequentialist theory knowledge. However, there does not appear to be explicit specified kinds of causal studies indicating that children are influenced by some kind of ethical theory knowledge. However, in a correlational study pertaining to folk consequentialism, Korean children appear to justify their moral beliefs based on social status, social roles, and the overall welfare of society as well; justifications that are not commonly observed in children in the United States.[17] Although we may infer that theory-theory folk consequentialist concepts are acquired by many individuals, to make a stronger case for concept acquisition, it must be admitted that explicit specified kinds of causal studies need to be run on children and adolescents regarding whether they are influenced in the appropriate qualified way by some kind of ethical theory knowledge when making moral judgments.

---

[17] Song, J. Smetana, Kim, 1987.

4.3 Prinz's Psychopath Argument for Epistemic Emotionism.

Before arguing that the emotion theory prima facie is a viable view of moral concepts based on causal evidence, we will criticize Prinz's psychopaths-based argument for the emotion theory. The objections to Prinz's contention open the pathway to show in the next section of this chapter that using appropriate causal evidence rather than studies on psychopaths is the more successful way to argue for the emotion theory. We now will examine what I take to be the strongest argument by Prinz that supports his epistemic emotionism, where as discussed in the second chapter, this view states that all moral concepts are constituted by sentiments and at times by emotions. Sentiments are dispositions to feel emotions. The criticisms against this argument can be taken to help show that using causal studies in a qualified manner perhaps may be the best empirical means for supporting the viability of an emotion-based theory for moral concepts.[18]

---

[18] One may think that Jonathan Haidt offers an empirical argument that may be read as claiming that all moral concepts are constituted by emotions. One may think that his Social Intuitionist Model (SIM) claims that moral judgments are influenced by emotions and that all reasoning is post hoc. Thus, all moral concepts are constituted by emotions. However, this is an incorrect understanding of his view in that he believes that emotions only causally influence judgment. They do not constitute judgment. He fails to see that a specified kind of causation can lead to concept constitution claims (this also holds for Nichols). Furthermore, he allows for reasoning to influence judgment in some cases. Moreover, he claims that judgments are mostly influenced by intuitions; intuitions which may or may not be generated by emotions. "While some critics erroneously reduce the SIM to the claim that moral reasoning doesn't happen or doesn't matter, we point out that the SIM says that reasoning happens between people quite often, and within individuals occasionally. Furthermore, the SIM is not about "cognition" and "emotion"; it is about two kinds of cognition: fast intuition (which is sometimes but not always a part of an emotional response) and slow reasoning." Jonathan Haidt and Fredrik Bjorklund, "Social Intuitionists Answer Six Questions about Moral Psychology," in *Moral Psychology, Vol 2*, ed. by Walter Sinnott-Armstrong, (Cambridge, MA: The MIT Press), 2008, p. 200. Jonathan Haidt, "The

Prinz provides an empirical-based argument for his epistemic emotionism view based on the recent studies and interest in psychopaths in moral psychology. Prinz turns to moral psychology studies on psychopathic individuals who have severe deficiencies in the emotions. Psychopaths appear to make sincere moral judgments, but they seem to lack the emotion-backed motivation to act upon their judgments. Psychopaths are known for their deficiencies in affect as well as their participation in heinous crimes. At face value, they appear intelligent and able to make sincere moral judgments, but they fail to act upon such judgments and may lack the moral motivation to so act based on their lack of emotions.[19] Thus, at first glance, it does seem like psychopaths do provide evidence against epistemic emotionism. To note, psychopaths differ from frontotemporal dementia patients in that although they are both emotionally shallow, psychopaths display more deliberate deceitfulness and have the need for constant stimulation.[20]

However, Prinz as well as James Blair and Shaun Nichols have pointed out that psychopaths apparently cannot distinguish between supposed moral versus conventional norms.[21] Thus, it does not appear to be the case that psychopaths make sincere moral

---

emotional dog and its rational tail: A social intuitionist approach to moral judgment," *Psychological Review, 108*: (2001), 814-834.

[19] H.M. Cleckley, *The Mask of Sanity: An Attempt to Reinterpret the So-Called Psychopathic Personality*, (St. Louis, MO: The C.V. Mosby Company), 1941.

[20] Mario Mendez, "What frontotemporal dementia reveals about the neurobiological basis of morality," *Medical Hypothesis* 67: (2006), 411-418.

[21] I use the terms 'apparent' and 'supposed' because I will later present evidence that psychopaths can make the moral/conventional distinction and that Prinz's moral/conventional

judgments. If they are incapable of making sincere moral judgments, then they do not provide empirical evidence against epistemic emotionism. To note, we will shortly discuss how Prinz also uses the psychopath studies to make a positive argument for his epistemic emotionism. Numerous contemporary philosophers have attempted to make the distinction between the ethical and the conventional. For example, in the edited volume *The Definition of Morality*, Alasdair MacIntyre, Elizabeth Anscombe, Philippa Foot, Peter Strawson, and others attempt to draw such a distinction a priori.[22] However, all the philosophers reach different conclusions and fail to reach any consensus. On the other hand, recently psychologists have begun positing that there supposedly are respective properties to the moral and the conventional domains that may distinguish the two and that also are relevant to scientific induction and explanation. Such properties are psychologically real and psychologically important. In other words, psychologists have argued by drawing on empirical data rather than by relying on a priori intuitions that morality is a purported natural kind and has a set of related properties in which scientific inductive generalizations may be drawn.

---

distinction is not legitimate. R. Blair, "A cognitive developmental approach to morality: Investigating the psychopath," *Cognition* 57: (1995), 1-29. Shaun Nichols, *Sentimental Rules*, (Oxford: Oxford University Press, 2004).

[22] In the final chapter of this dissertation, we will attack the method of conceptual analysis that is commonly used to draw an a priori distinction between the moral and the conventional. Ed. by G. Wallace and A. Walker, *The Definition of Morality*, (London: Methuen & Co., 1970).

Prinz slightly strays from the work done by Turiel, Smetana, Nucci, and other psychologists on the supposed moral/conventional task by providing a more narrow list of what the properties are for moral versus conventional violations and norms.[23] Prinz may be read as claiming that that there is a robust scientifically significant consensus on the ostensible moral/conventional task in that cross-culturally, harm norms such as those against killing, stealing, and lying involve a moral rather than conventional response pattern that is characterized by the property of being a grounding norm or being justified in terms of a grounding norm. Here a grounding norm is a norm that does not require any further explanation or justification. For example, norms against killing and stealing supposedly do not require any further justification. They are foundational norms in that normal folk, for example, purportedly do not ask for and do not require a further justification as to why killing in cold blood is wrong. Harm norms fall within the domain of morality because they supposedly are grounding norms or justified in terms of grounding norms, and they are not necessarily justified by appeal to local customs. As we can see, this grounding norm property is intimately linked to the property of authority independence or scope objectivity in applying equally across location and time, where having scope objectivity means that the rule in question has the broad scope of applying

---

[23] Unlike the psychologists, Prinz does not maintain that morality and the conventional are supposed homeostatically clustered natural kinds. However, the arguments presented here against Prinz's moral/conventional task can also be used against a homeostatically clustered natural kinds theory of the moral and conventional. The standard list of properties for harm transgressions in the homeostatically clustered natural kind moral/conventional task literature is: scope objectivity, authority independence, justification by appeal to rights or harm, serious harm or injustice.

across cultures and time. For, if a harm norm has the grounding norm property and is not justified by appeal to custom, then it will be taken to have scope objectivity rather than being limited to a particular culture or period of time.

On the other hand, as a different scientific generalization, non-harm norms such as those against failing to take off one's hat while eating, drinking soup from a bowl, and speaking in class without raising one's hand purportedly evoke the conventional response pattern in children and adults of being justified by appeal to customs. For example, we supposedly will say that the reason why we take off our hats when eating is due to the fact that this is the way we do things around here. Moreover, this generalization purportedly holds across cultures. Notice the intimate link between this property and that of authority dependence or lack of scope objectivity. For, if we justify a non-harm norm by way of custom, then we will take the norm to be authority dependent or lacking in scope objectivity.

All in all, Prinz writes:

> If I am right (and the literature on the moral/conventional distinction bears this out), people do not ordinarily appeal to local customs when justifying harm norms. They don't say that it's wrong to whip [humans] because people don't whip around here. In contrast, if you asked people why it's wrong to wear a hat while eating, they would be more likely to mention local customs.[24]

However, psychopaths apparently treat both harm and non-harm norms as authority independent. For example, even if the teacher says that one may speak out in class, psychopaths still replied that it is wrong to speak out in class. Blair hypothesizes

---

[24] Prinz, *Emotional,* p. 128.

that since he examined imprisoned psychopaths, such subjects have an interest in pleasing their interviewers in the hopes of receiving an early release.[25] Thus, psychopaths apparently claim that all rules, moral or conventional, are inviolable at all times, which stands in opposition to normal prisoners who are able to successfully complete the supposed moral/conventional task. Psychopaths appear to be able to get a few of the questions right on the task, but they do significantly worse than normal prisoners on the task. Prisoners who truly understand the ostensible moral/conventional distinction and desire to show their moral competence supposedly will be able to draw the distinction. Thus, Blair claims that psychopaths really cannot distinguish between the supposed moral and the conventional because they would have done so if asked. Rather, it apparently is the case that they really view morality as equivalent to conventional norms and merely pay lip service to morality. This hypothesis supposedly is further backed by Blair's ostensible moral/conventional task study on children with psychopathic tendencies.[26] Due to these findings, Prinz claims that it is not the case that psychopaths have proper moral concepts and can make sincere moral judgments. Thus, Prinz argues that psychopaths cannot be used as evidential support against epistemic emotionism.

Prinz takes this a step further and adds that there appears to be a necessary link between emotions and moral competence. Psychopaths have an emotional deficit which co-occurs with the fact that they supposedly cannot make the apparent

---

[25] Blair, ibid.

[26] R. Blair, "Moral Reasoning and the Child with Psychopathic Tendencies," *Personality and Individual Differences* 25: (1997), 731-9.

moral/conventional distinction. This suggests that one must have moral sentiments in order to comprehend morality. In other words, for Prinz, to have a correct moral conception in order to make a sincere moral judgment, one's concept must be essentially related to emotions. Thus, Prinz believes that the case of psychopaths does not provide evidence against epistemic emotionism, but rather, it actually supports the case for epistemic emotionism.

At this point, I will present arguments against Prinz's naturalistically supported view of epistemic emotionism. First, psychopaths apparently are not completely bereft of sentiments, so it may not be the case that a deficiency of sentiments is responsible for their apparent inability to perform well on the supposed moral/conventional task. For example, although at a more reduced level as compared to non-psychopathic subjects, psychopaths can still show significant galvanic skin responses to pictures of individuals who are under distress and crying.[27] This is suggestive of a reduced but still present empathy response. It may be the case that the sentiments psychopaths apparently do have are all that is needed to make the supposed distinction or it may be the case that absolutely no sentiments are required at all to complete the task, and some other mental deficiency is responsible for their supposed inability to distinguish between the purported moral and conventional. Research on psychopaths is still in its developing stages, and there are several studies indicating that psychopaths have some cognitive deficiencies as

---

[27]J. Blair, J. Morris, C. Frith, et al.,"Dissociable neural responses to facial expressions of sadness and anger,"*Brain*, 122, 1999, 883 -893.

well.[28] For example, in one study, Ethan Gorenstein has shown that psychopaths do worse than controls in the Wisconsin Card Sorting Task (WCST) and the Sequential Matching Memory Task (SMMT).[29] The WCST asks participants to sort or categorize a stack of cards into various piles based on certain qualities that the pictures on the cards contain. The WCST then tests to see how well subjects are able to flexibly adapt to the categorization process once the rules of categorization have changed. The SMMT is a cognitive memory test in which participants are shown a series of cards one by one that either have a plus or minus sign on them. Subjects are tested as to whether they can recall the sign of the card two cards previous to the one currently displayed. While these studies suggest there are cognitive deficiencies in psychopaths related to categorization and memory, I contend that they also open up the possibility that psychopaths may also have deficiencies in moral categorization and moral reasoning. Without any evidence to the contrary, this possibility remains open, and at this point in the study of psychopaths, I argue that Prinz's reliance on psychopaths to contend for epistemic emotionism may be put in some doubt, and it is too premature to absolutely conclude that blunted emotions are responsible for psychopaths having apparent problems with the purported task. It is possible that the sentiments they supposedly do have may be enough to adequately

---

[28] See D. Schalling and A. Rosen, "Porteus maze differences between psychopathic and Non-psychopathic criminals," *British Journal of Social and Clinical Psychology* 7: (1968), pp. 224-8. Joshua Zeier, Arielle Baskin-Sommers, Kristina Racer, and Joseph Newman, "Cognitive Control Deficits Associated with Antisocial Personality Disorder and Psychopathy," *Personality Disorder: Theory, Research, and Treatment*, forthcoming, advanced on-line publication (July 4, 2011).

[29] E. E. Gorenstein, "Frontal Lobe Functions in Psychopaths," *Journal of Abnormal Psychology*, 91: (1982), p. 368-379.

perform the ostensible task or that absolutely no sentiments are required at all to complete the purported task, and it may be deficits in moral cognitive functions that really are responsible for their apparent failure in the supposed task.

Second, since psychopaths can still experience emotions on a limited basis, they are still conceivably disposed to feel emotions, although such a disposition does not manifest itself occurrently on a normal basis. Recall that sentiments are dispositions to feel emotions. For example, psychopaths at times can feel the moral emotion of anger when they have been rejected by other people, especially those close to them such as by family and friends.[30] They can also at times have a hypersensitivity to anger when others refuse to cooperate with them.[31] Death of a loved one can also spur psychopaths to

---

[30] Willem Martens, "Antisocial and psychopathic personality disorders: causes, course and remission," *International Journal of Offender Therapy and Comparative Criminology*, 44(4): (2000), 406-430.

[31] While some may perceive the disposition to feel anger to be less important than the disposition to feel the pain and suffering of others when making moral judgments, I list this study here (along with other studies suggesting that psychopaths also may have the disposition to feel the pain and suffering of others) because the disposition to feel anger is still thought to influence moral judgments even though it may play a lesser role as compared to other dispositions. The sentiment of anger is still commonly considered to be a moral sentiment. For a study on the importance of anger to moral judgments, see E. Fehr and S. Gachter, "Altruistic punishment in humans," *Nature* 415: (2002), 137-140. For studies indicating that psychopaths have a hypersensitivity to anger, see H. Kohut, *The analysis of the self*, (New York: International Universities Press), 1971. O. Kernberg, "A Psychoanalytic classification of character pathology," *Journal of the American Psychoanalytical Association* 18: 1970, pp. 800-822. O. Kernberg, "Early ego integration and object relations," *Annals of the New York Academy of Sciences* 193: 1972, pp. 233-247. O. Kernberg, *Aggression in personality disorders and perversions*, (New Haven: Yale University Press), 1992.

experience guilt for their previous criminal activities.[32] Moreover, psychopaths can feel the pain and suffering of others in that they can feel empathy and sympathy towards those with whom they have a long-lasting and warm relationship.[33] Recall that we have previously discussed how psychopaths still have an empathy response when shown pictures of individuals who are in distress or crying. Experiencing a string of good luck or occupational success can also at times elicit pro-social emotions such as caring.[34] Skin conductance response and cardiac deceleration tests have suggested that psychopaths have delayed emotional reactions to unpleasant emotionally charged pictures.[35] Moreover, blood pressure, pulse, and self-report tests indicate that psychopaths feel anger in the performer evaluator paradigm.[36] In this experiment, participants witness a person adequately performing a given task then being unjustly criticized on his performance by a confederate or actor. Indeed, the University of Wisconsin psychologist Joseph Newman and others controversially hypothesize that although psychopaths generally do have blunted emotions as compared to normal subjects, at numerous times psychopaths do

---

[32] Willem Martens, *Psychopathy and Maturation*, PhD-thesis, Tilburg University, The Netherlands, (Maastricht: Shaker Publishing), 1997. Willem Martens, "Emotional Capacities and sensitivity in Psychopaths," *Dynamical Psychology*: 2003.
[33] Ibid.
[34] ibid.
[35] Steven Sutton, Jennifer Vitale, and Joseph Newman, "Emotion Among Women With Psychopathy During Picture Perception," *Journal of Abnormal Psychology* 111: 2002, pp. 610-619.
[36] Brian Lee Steuerwald, *Anger Following Provocation in Individuals with Psychopathic Characteristics*, (Greensboro, NC: University of North Carolina Greensboro Press), 1996.

experience emotions in social interpersonal settings, but they generally do not selectively pay attention to their emotions.[37]

Therefore, it appears that psychopaths may have sentiments or dispositions to feel emotions in relations to certain moral emotions such as anger and empathy.[38] Although this may be the case, whether psychopaths truly have any sentiments or not may still be a controversial issue in the psychopaths literature given that there are many studies also suggesting that at times they have blunted affect in relation to emotions such as anger and empathy.[39] However, we may still embrace this controversial topic and use it to provide a new objection to Prinz. For, at minimum, at this stage it seems that it is at least highly controversial whether psychopaths have sentiments or not. Perhaps more studies are needed to make a decisive claim on this matter. Thus, at minimum and assuming that psychopaths cannot successfully complete the supposed moral/conventional task,[40] I contend that Prinz's initial epistemic emotionism thesis cannot be adequately justified based on psychopath studies since we are not sure whether psychopaths are disposed to feel emotions or not. It may be, for example, an inability to pay attention to one's

---

[37] A. Baskin-Sommers, J. Curtin, J. Newman, "Specifying the attentional selection that moderates the fearlessness of psychopathic offenders," *Psychological Science 22(2): 2011, pp.* 226–234.

[38] Empathy is commonly understood as the feeling of another's emotions and realizing that one's feelings result from theirs. Insofar as it involves the feeling of an emotion, empathy is typically understood as being an emotion.

[39] One issue in particular that makes this a controversial topic is that further psychological theorizing will be required to distinguish between a psychopaths only sometimes being associated with experiencing emotions versus psychopaths actually having a disposition to feel emotions which is not always realized.

[40] It will later be shown that psychopaths can successfully complete the moral/conventional task.

emotions that psychopaths are disposed to feeling that explains psychopaths' apparent inability to perform the supposed moral/conventional task, or perhaps it may be a lack of emotional dispositions that is responsible. Since we are unsure of this matter, I argue that Prinz cannot conclude with the requisite certainty that all moral concepts are constituted by sentiments; a conclusion he draws that is based in part on the questionable fact that psychopaths lack moral sentiments.

Third, a recent study has shown that psychopaths can successfully complete the supposed moral/conventional task. In order to eliminate the possible design flaw in previous experiments that incarcerated psychopaths can really distinguish between the purported moral and the conventional but they do not do so because they want to impress the parole board by inflating conventional harms to that of moral ones, Aharoni, et al. gave the supposed moral conventional task to groups of psychopathic inmates and non-psychopathic inmates.[41] However, before answering the sixteen questions, subjects were told in advance that eight of the acts were pre-rated to be morally wrong, or, in other words, wrong even if there were no conventions or laws against them. Moreover, they were told that eight were pre-rated to be conventionally wrong, or, in other words, not wrong if there were no conventions or laws against them. In this fashion, our experimenters attempt to eliminate the possibility that psychopaths are attempting to impress others with their answers. Participants were asked to label each violation

---

[41] E. Aharoni, W. Sinnott-Armstrong, and K. Kiehl, "Can Psychopathic Offenders Discern Moral Wrongs? A New Look at the Moral/Conventional Distinction," *Journal of Abnormal Psychology, forthcoming.*

according to whether or not the act would be wrong if there were no conventions or laws against it. Notice that Prinz's supposed moral/conventional distinction predicts that if psychopaths cannot really distinguish between the moral and the conventional, then they should do rather poorly on this test since they do not understand that harm norms are not justified by appeal to custom but non-harm norms are so justified. The results of the experiment were that psychopaths indeed were able to distinguish between purported moral and conventional norms comparable to the non-psychopathic inmates. Aharoni et al. ran their experiments in order to challenge the fact that psychopaths cannot successfully complete the moral/conventional task. Hence, we now can use their tests and apply them to Prinz and his claim on moral concepts. Given that there is evidence that psychopaths do actually properly make the supposed moral/conventional distinction, it appears to not be the case that all moral concepts must be constituted by sentiments and at times by emotions based on psychopath studies.[42] While surely more studies must be

---

[42] One may object to this study in that the number of conventional versus moral harms is already given to subjects and this may provide weak cues to make it easier for participants to make the supposed distinctions. This may be analogous to how color-blind persons can guess which colors are which based on the weak cues of lightness and contextual appropriateness. However, one must ask what the weak cues may be in the apparent moral/conventional case. If there are cues, they allow for psychopaths to do just as well as non-psychopathic inmates. I see no other possibility but that it must be some features of the situations that allow psychopaths to successfully complete the ostensible task. However, if this is the case, then emotionally blunted psychopaths must be relying on some kind of conscious or subconscious knowledge of a purported moral/conventional distinction as it is certain features of the situation that are supposed to allow normal subjects to differentiate between the moral and the conventional. Unlike the color-blind case, there does not appear to be some weak or intermediate cue, such as contextual appropriateness, in the supposed moral/conventional case that allows one to tell the

conducted in order to fully determine whether psychopaths can make the supposed moral/conventional distinction, this study at hand places enough doubt on previous conclusions that psychopaths cannot make the purported moral/conventional distinction, that Prinz's reliance on such previous conclusions in order to support epistemic emotionism is unjustified.

Fourth, I will now attack Prinz's moral/conventional task. If successful, this will show that the premise that 'psychopaths cannot distinguish between the moral and the conventional' may be significantly doubted since Prinz's moral/conventional distinction is not legitimate. Given that this is the case, the premise cannot be said to support epistemic emotionism. It must be stated that there may not be enough evidence to completely discount Prinz's moral/conventional task, but what will be attempted here is to demonstrate that there is enough evidence to cast significant doubt on it; a significant

---

difference between the moral and the conventional without actually knowing the difference between the moral and the conventional.

Telling subjects how many moral and conventional violations there are is beneficial in that it eliminates the possibility that psychopaths claim that all violations are moral in order to impress the parole board. Nevertheless, this may create problems in that it makes it easier for subjects to perform the task given that they know the number of situations that belong in each category. While it may make it easier in this respect, if psychopaths really do not know how to make the supposed distinction, then we should still expect them to do worse than non-psychopathic inmates. However, they do just as well as non-psychopathic inmates, and this must be due to the fact that it is based on certain features of the situations that allows them to successfully complete the study. Yet if this is the case, then it appears that they really do know how to distinguish between the purported moral and the conventional because it is thought that such a distinction is usually made based on situational features.

doubt that may also be placed on any larger conclusions such as epistemic emotionism that in this particular argument crucially relies on the legitimacy of this task.

A host of philosophers and psychologists have run experiments in order to explicitly or implicitly challenge the moral/conventional task. For example, Jonathan Haidt, et al. ran an experiment where they discovered that lower socio-economic subjects in the United States and in Brazil judged non-harmful transgressions to be authority independent and to have scope objectivity.[43] For example, when such participants were asked to make judgments on a man who privately has sexual intercourse with a dead chicken before cooking and eating it or on a woman who privately cleans her toilet bowl with the national flag, they answered that these cases are universally wrong even though they perceived such transgressions to be harmless. On the other hand, higher socio-economic participants did not hold such cases to be universally wrong. Thus, this study indicates that perceived harmless transgressions that should elicit an opposing conventional response pattern can actually elicit the moral response pattern of authority independence and scope objectivity. Hence, whether or not one perceives certain harmless transgressions to have scope objectivity and authority independence may be relative to one's socio-economic status. This study suggests that justifications that do not appeal to local custom are being used in non-harm cases. Additionally, Korean children also claimed that the non-harmful transgression of not appropriately greeting one's elders

---

[43] J. Haidt, S. Koller, M. Dias, "Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog?" *Journal of Personality and Social Psychology* 65: (1993), 613-28.

as having the moral response pattern properties of scope objectivity and authority independence.[44]

Shaun Nichols has run studies on disgusting and non-disgusting violations of etiquette norms.[45] For example, a disgusting non-harm transgression is when a dinner guest snorts then spits into his water glass and drinks from it. A non-disgusting non-harm violation is when a dinner guest drinks soup directly out of his bowl. He found that children believed disgusting transgressions to be authority independent and as having scope objectivity. Meanwhile, adults also took such disgust violations to be authority independent. The above studies provide evidence against the scientific generalization that cross-culturally, violations that do not involve stereotypical harm transgressions evoke the conventional response pattern.

Moving in the opposite direction where we will be discussing participants' reactions to violations that do involve prototypical harm transgressions, Hagop Sarkissian and his collaborators have suggested in a cross-cultural study with participants in the U.S. and Singapore that clear harm violations can elicit aspects of a conventional response pattern.[46] For example, participants are given scenarios in which an individual kills

---

[44] M. Song, J. Smetana, S. Kim, ibid.

[45] Shaun Nichols, "Norms with feeling: Toward a psychological account of moral judgment," *Cognition* 84: (2002), 223-236. Nichols, *Sentimental Rules*.

[46] The findings from this study contravene the Goodwin and Darley study that suggests that there is a general underlying objectivity to ordinary moral discourse (Goodwin & Darley, ibid.). For, the interlocutors in the Goodwin and Darley studies were all in the same culture, and subjects may have assumed that the interlocutors have generally the same values even though they may reach different conclusion on certain ethical matters. However, once individuals from different

random innocent bystanders. In each case, the participant is told that one's classmate believes such random acts of violence are morally wrong but a member of a hypothetical alien culture holds the opposite view. Subjects are then asked about these conflicting judgments and whether they believe it is the case that only one of the disputants may be right or that both of them can be right. If participants believe that both of them can be right, then this shows evidence that they believe in an underlying cultural moral relativism concerning certain instances when interlocutors have differing moral views, and that such harm transgressions in question do not have scope objectivity. This study found that a majority of participants thought that both disputants could be right in a number of cases and that they did not believe that such harm transgressions are characterized by scope objectivity. This study implies that beliefs about local customs factor into judgments about paradigmatic harm norms.[47]

---

cultures with a significantly different set of values is introduced as in the Sarkissian, et al. experiments, an underlying relativism is drawn out from subjects. H. Sarkissian, J. Park, D. Tien, J. C. Wright, J. Knobe, "Folk Moral Relativism," *Mind & Language* 26: (2011), pp. 482-505.

[47] Daniel Kelly and company have run studies attempting to challenge whether harm violations are characterized by the moral response pattern. For example, they found that participants believed a ship captain whipping his derelict and drunken sailor 300 years ago lacks authority independence. However, I do not put forth this study as evidence against the supposed moral/conventional task due to objections given by Prinz. For example, Prinz argues that subjects state that it is permissible for the historical captain to whip his sailor because there are historical factors that allow for the moral permissibility of the act – e.g., that sailors in the past were more prone to unruly behavior and mutiny than the professional sailors of today's age. To note, through verbal communication, Stich has said that he now has the data from running moral/conventional task studies in twelve different cultures. The results suggest that Prinz's moral/conventional distinction is not legitimate. Kelly, et al., ibid. Prinz, *Emotional*, 127-8.

Likewise, Nichols ran a series of experiments where in one test, subjects are asked to imagine two people from different cultures debating whether it is permissible to hit another person just for the fun of it.[48] One person is from a culture where it is not permissible while the other interlocutor is from a culture in which it is permissible. Nichols found that 66% of the subjects believed that the moral transgression of hitting others for fun lacks scope objectivity.[49] This study also implies that beliefs about local customs factor into judgments about paradigmatic harm norms. Moreover, in a different study in Israel, traditional Arab children differed from urban Jewish children in that their justification of why harm transgressions are wrong is based on an appeal to customs.[50] While many more studies need to be conducted in order to make a complete case, these above studies regarding subjects' reactions to harm transgressions begin to suggest that cross-culturally, such transgressions do not evoke the typical moral response pattern.

---

[48] Shaun Nichols, "After Objectivity: An Empirical Study of Moral Judgment," *Philosophical Psychology* 17: (2004), 5-28.

[49] Several of the other studies in Nichols' experiments contrarily showed that a majority of subjects were objectivists, but there was still a significant group of non-objectivists in these cases. For example, 17 of the 40 participants in one of the experiments were non-objectivists while 23 were objectivists. However, in the Sarkissian, et al. studies, they ran six experiments in which they received consistent results where most subjects were non-objectivists. Perhaps one explanation for the consistency in the Sarkissian, et al. test results as opposed to the Nichols experiments is that Sarkissian and company gave a one page detailed description of the alien culture while Nichols did not and merely stated that the two interlocuters are from different cultures without providing further details. The greater detail perhaps drove home the point that the debaters were from different cultures with vastly different ways of life. This factor may have led to the consistency in Sarkissian, et al.'s various experiments where a majority of participants were non-objectivists in all six studies.

[50] M. Nisan, "Moral norms and social conventions: A cross-cultural comparison," *Developmental Psychology* , 23 : (1987), 719 – 725 .

While it is also admitted that more work needs to be done overall on the supposed moral/conventional task to reach a decisive conclusion whether Prinz's moral/conventional distinction holds, what has been attempted here is to show that enough doubt can be placed on this task that any concepts conclusion that is based on Prinz's moral/conventional distinction – or for that matter, any such universal distinction, however based – is unwarranted barring further evidence. Given that this is the case along with the other three previously given arguments, we cannot draw a strong conclusion from Prinz's psychopaths-based argument that all moral concepts are dispositionally and at times constitutionally related to emotions.

4.4 The Arguments for the Emotion Theory

Now that we have seen how Prinz's psychopaths-based argument for the emotion theory does not work, we will explore how appropriate causal studies in moral psychology can be used to argue for the emotion theory. In order to understand experiments that may be used to make a prima facie viability claim for the emotion theory for moral concepts, a basic understanding of the trolley problems is required.[51] In the trolley cases, imagine there is a *lever case* in which a train is about to run over five people, but you may pull a lever to divert the train on a side track in order to save the five

---

[51]J. Greene, R.B. Sommerville, L.E. Nystrom, J.M. Darley, and J.D. Cohen, "An fmri investigation of emotional engagement in moral judgment," *Science* 293, (2001), 2105-8. J. Greene, L.E. Nystrom, A.D. Engell, J.M. Darley, and J.D. Cohen, "The Neural Bases of Cognitive conflict and Control in Moral Judgment," Neuron 44: 2004, 387-400.

people. However, it just so happens that there is one person who is on the side track who will be killed if you pull the lever to save the five. 89% of participants agree that one ought to pull the lever in this scenario. Now, in the *footbridge case*, imagine that the same train is heading towards the same five people, but now you are on a footbridge that hovers over the track. In your company is a heavyset man whose body mass is large enough to stop the train if you kill him by pushing him over the bridge on to the track in order to save the five. In this case, 89% of participants state that one should not push the heavyset man. Greene interprets the trolley cases as a matter between consequentialism and deontology.[52] Studies have shown that the lever case is one in which most participants make supposed folk consequentialist judgments in which it is permissible to pull the lever in order to save a greater number of lives while the footbridge case is one in which most subjects make supposed Kantian deontological judgments in which the heavyset man has an individual right not to be pushed regardless of the greater consequences that will arise in pushing him. In these trolley studies, when stating that subjects make deontological judgments, what is meant is that they draw conclusions that

---

[52] Some interpret this set of trolley problems as a matter of the Doctrine of Double Effect. The Doctrine states that at times it is permissible to harm others as a side effect of attaining a good end, but it is not permissible to harm others as a means of attaining a good end. However, as we will later see, I interpret the set of trolley problems as primarily being between folk consequentialism and emotions. It is doubtful that the Doctrine is being used in the lever and footbridge circumstances because if this is the case, then the same brain region(s) should be in play for the two trolley problems in the set since the Doctrine supposedly is at work in both these cases. However, this is not the case. The dorsolateral prefrontal cortex is primarily active for the lever trolley case, but the ventromedial prefrontal cortex region is primarily active for the footbridge case. This provides suggestive yet not quite conclusive evidence that the Doctrine is not being used in these particular trolley problems.

would be thought to be made by deontological normative theory. It is not meant to necessarily mean that deontological reasoning is actually used to arrive upon a moral conclusion. However, concerning folk consequentialist reasoning, it will later be argued with additional evidence that such reasoning is being used by normal subjects in the lever case. Now, as with Greene, et al.'s, cognitive load study, taken by itself, it is not decisive that subjects are truly making folk consequentialist or, for that matter, deontological judgments, but as a matter of convenience, we shall abide by these classifications for the two different types of judgments for now.

Mendez, et al. have run trolley studies on patients with frontotemporal dementia which involves a deterioration of the ventro-medial prefrontal cortex (VMPFC) and anterior temporal emotion areas of the brain.[53] Although they are thought to have normal reasoning processes, these patients have been found to not only show diminished concern for others but they also exhibit emotional blunting or severely diminished emotions such as having a lack of empathy. Due to emotional blunting, such patients are prone to committing transgressions such as stealing and inducing violent physical harm to others. They do not respond to the needs of others nor acknowledge them, and they also display inappropriate sexual behavior such as pedophilia. Mendez, et al.'s study has been

---

[53] M.F. Mendez, E. Anderson, & J.S. Shapria, "An investigation of moral judgment in frontotemporal dementia," *Cognitive and Behavioral Neurology*, 18: 4 (2005), 193-7.

successfully replicated by Koenigs & Young, et al. and Ciaramelli, et al. on patients that also have lesions in the VMPFC.[54]

Relatedly, in *Descartes' Error*, Antonio Damasio has similarly concluded that patients with lesions to the VMPFC have intact reasoning capacities but are considered to have deficiencies in affect.[55] For instance, Damasio describes the case of one of his patients who scores normally and at times above normal on intelligence and reasoning tests, but exhibits diminished emotions. When discussing the many tragic hardships of his life, the subject displays an unusual emotional detachment from such events with no sign of frustration or sadness. Moreover, when shown visually stimulating and emotionally charged pictures of people drowning and individuals being in gory accidents, the patient showed no emotional response and remained emotionally neutral. This patient's everyday life was generally characterized as one of disaffection, which is opposite from the time period in his life before he suffered damage to the VMPFC.

In Mendez, et al.'s work, patients with frontotemporal dementia were given the lever case and footbridge case. In the lever case, patients' responses matched that of normal subjects. For example, in the lever scenario most patients stated that it is permissible to pull the lever. However, in the footbridge case, patients significantly

---

[54] M. Koenigs, L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser, A. Damasio, "Damage to the prefrontal cortex increases utilitarian moral judgments," *Nature* 446: (2007), 908-911. E. Ciaramelli, M. Muccioli, E. Ladavas, and G. di Pellgrino, "Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex," *Social Cognitive and Affective Neuroscience* 2: (2007), 84-92.

[55] Antonio Damasio, *Descartes' Error*, (New York: Penguin Books), 1994.

diverged from the answers from normal subjects by allowing for the pushing of the heavyset man onto the track. With blunted emotions, frontotemporal dementia patients were more inclined to make apparent folk consequentialist judgments in this situation, while those with normal emotions tended to believe that pushing the heavyset man is not permissible. This study provides causal evidence that in certain cases such as the lever case, some kind of folk consequentialist reasoning influences moral judgment. After all, the patients who have blunted emotions made the same judgments as normal subjects in such cases, and normal subjects displayed activation in the dorsolateral prefrontal cortex as in Greene's cognitive load study.[56] Along with the evidence presented in the theory-theory section of this chapter concerning Greene's cognitive load study, neuroimaging studies, and cross-cultural evidence for a Kohlbergian folk consequentialism, this suggests that normal subjects prima facie may primarily be using a Kohlbergian folk consequentialism knowledge in making such decisions.[57] Here, the possibility is left open that emotions can still have some influence on judgment, but regardless of whether this is so, from this study we still may infer that a Kohlbergian folk consequentialism plays a prima facie role.

However, since the emotionally blunted patients diverged in judgments from normal participants in the footbridge case by making apparent folk consequentialist rather

---

[56] Greene, ibid.

[57] In *The Emotional Construction of Morals*, Prinz states that it is really the emotions driving judgments in the lever case since there is a small activation of emotion regions of the brain when subjects make such categorizations. However, Prinz mistakenly uses only correlational fmri data to draw a causal conclusion. Prinz, *The Emotional Construction of Morals*.

than deontological categorizations, this and other replication studies suggest that emotions do play a causal role in making judgments for normal subjects in the footbridge case. In that emotionally blunted patients did not make the same judgments as normal subjects did in the footbridge case, this suggests that normal participants are influenced by emotions when making such categorizations. Likewise, some kind of reasoning may also play a role alongside emotions in footbridge case categorizations for normal subjects, but whether or not this is the case, we still may infer that emotions play a prima facie specified causal role, and for our purposes, this is important. This replicated study provides prima facie evidence that the emotion theory is a viable view of moral concepts.[58] To note, the retraction, inferential base, and informational value tests have

---

[58] There is an apparently contradictory ultimatum game study run by Koenigs and Daniel Tranel. In typical ultimatum games, two players can split a sum of money, where the first player makes an offer, and the second can either accept or reject the offer. For example, a fair offer would be to split ten dollars equally, and an unfair offer would be an $8/$2 split. A convergence of studies suggest that rejections of unjust offers are influenced by emotional reactions of frustration and anger. However, Koenigs and Tramel discovered that VMPFC patients had a higher rejection rate of the most unfair offers ($7/$3, $8/$2, $9/$1) as compared to normal subjects. They conclude that VMPFC subjects experience anger and frustration which causes them to have a higher rejection rate of unfair offers and that damage to the VMPFC disrupts the ability to regulate the emotions such that one may harness one's anger and accept unfair offers in order to "rationally" attain at least some financial gain for one' self-interest. This particular study apparently contradicts the previous VMPFC studies related to the trolley problems in that VMPFC subjects, despite generally having blunted emotions, are known to experience emotions such as anger and irritability in certain situations, and such emotions influence judgment. For example, even Damasio notes how VMPFC patients, in rare instances, may lose their temper (*Descartes'* 45). However, when they do so, the outburst is swift, and patients quickly return to their calm detached states without bearing any grudges. This seems to put into question or to contradict the fact that in the lever and footbridge case it is thought that for VMPFC patients, emotions play little to no role in affecting such moral categorizations.

not been run on the normal subjects in this study. However, remember that so long as the normal subjects do not appear to be experiencing some kind of mental error or a potential causal factor to such error, such as being hypnotized or being manipulated by experimenters by being placed in a dirty room that elicits a disgust affect, then a prima facie constitution claim may be made for the concepts of normal participants. The normal subjects in the above studies do not appear to be experiencing some kind of mental error or a potential causal factor to such error when undergoing the tests. Thus, a prima facie conclusion for the emotion theory may be drawn for normal subjects.

Concerning concept acquisition, we already know that there is a prima facie constitution claim for the viability of the emotion theory for adults based on categorization studies, so if this is the case, emotion theory moral concepts must have

---

However, Koenigs & Young, et al., note that the emotional effects of VMPFC damage are context dependent. VMPFC subjects may exhibit emotions such as anger and frustration when being directly and personally provoked such as in the ultimatum game when frustration is triggered due to unfair take-it-or-leave-it offers that also involves an element of self-interest and direct involvement. On the other hand, such subjects are generally characterized as having severely diminished social emotions related to such feelings as empathy and a concern for others, where such an emotionally detached state is manifested, *inter alia*, in hypothetical scenarios such as the trolley problems where the patient is not directly and personally involved in a real life situation. Thus, in the context of the VMPFC studies that include the trolley cases, VMPFC patients have blunted social emotions, and such a context is not a case in which the subject is personally and directly involved such that the emotions of anger and frustration may potentially manifest themselves and play a causal role. Therefore, we may infer that reasoning is the primary causal influence in moral categorizations by VMPFC patients in the lever and footbridge cases. Michael Koenigs and Daniel Tranel, "Irrational Economic Decision-Making after Ventromedial Prefrontal Damage: Evidence from the Ultimatum Game," *The Journal of Neuroscience* 27: (2007), 951-956. Koenigs & Young, ibid.

been acquired in some way. However, for concept acquisition, there are no studies explicitly showing that emotions in the specified kind of way influence rather than are correlated with children's moral judgments. However, it may be demonstrated here that there is a strong possibility that in future tests on children, the emotion theory may be shown to be viable. This strong possibility later will be inferred based on combining several facts such as that children do experience moral emotions and that developmental psychologists observe children to be influenced by emotions when making moral decisions.

While it is widely understood in folk psychology that children can experience emotions such as anger, there are studies showing that children do experience moral emotions. For instance, Nunner-Winkler and Sodian along with Nichols hypothesize that four year-olds do not comprehend the moral emotion of guilt yet, but such a comprehension is attained in later years.[59] However, this is controversial in that a number of studies have shown that children younger than four can comprehend guilt.[60]

---

[59] Nichols also uses this fact to mount a criticism against Allan Gibbard's metaethical view which in part states that an action is morally wrong if one accepts norms that obligates one to feel guilt when performing the action. Gibbard's view requires us to be able to form judgments about moral emotions such as guilt, but Nichols argues that four-year olds cannot form such judgments about guilt yet even though they have mastered the moral/convention task. In other words, such children moralize without the comprehension of guilt. Nichols, *Sentimental*. Allan Gibbard, *Wise Choices, Apt Feelings*, (Oxford: Clarendon Press), 1991.

[60] Michael Lewis, "Emotional Competence and Development," in *Improving Competence across the Lifespan,* ed. by D. Pushkar, W. M. Bukowski, A. E. Schwartzman, D.M. Stack and D. R. White, (New York: Plenum), 1998, 27-36. K. Barrett, C. Zahn-Waxler, and P. Cole, "Avoiders Versus Amenders: Implication for the Investigation of Guilt and Shame During Toddlerhood?" *Cognition and Emotion* 7: (1993), 481-505. Kagan, ibid.

For instance, Kochanska and company gave two-years-old children a doll that is said to be of special value to its owner.[61] The children are warned to be extremely careful with it, but the experiment is set up such that the doll eventually will end up breaking or ink will be spilled on it. The experimenters found that when this occurred, children showed signs of guilt by avoiding the experimenters gaze, squirming, or covering their faces with their hands. Such a study suggests that children may experience the moral emotion of guilt at a very young age.

Also, while not based on direct experimental causal evidence where emotions are shown to influence judgments in the appropriate qualified way, empathy is usually thought by developmental psychologists to be at work at times when adolescents reason about moral situations by placing themselves in others' shoes.[62] It is common for childhood psychologists to personally observe and then hypothesize that emotions at times influence and motivate certain judgments and actions in children and adolescents. For example, the Cornell psychologist Kenneth Barish claims that the emotion of shame is a powerful emotion that has more lasting power than most other emotions. [63] For example, if a child feels a strong sense of shame for seriously letting down her father on a moral matter, then this shame will tend to stick with the child, at times in suppressed

---

[61] G. Kochanska, J. Gross, M. Lin, and K. Nichols, "Guilt in Young Children: Development, Determinants, and Relations With A Broader System of Standards, *Child Development* 73: (2002), 461-82.

[62] Abigail Baird, "Adolescent Moral Reasoning," *Moral Psychology Vol. 3*, ed. by Walter Sinnott-Armstrong, (Cambridge, MA: The MIT Press), 2008.

[63] Kenneth Barish, *Pride and Joy*, (Oxford: Oxford University Press), 2012.

form, even through adolescent years, where it can in part later motivate her in certain relevant moral circumstances to do the right thing.

Despite the fact that all of the above studies do not directly and experimentally support concept acquisition for the emotion theory in children, such studies at least provide some optimism for a hypothesis on future studies on children and moral emotions that children will at times be influenced by emotions in the appropriate qualified way when making moral judgments. In this circumstance, recall that the appropriate qualified way for emotions is when 1) the emotion is used by default and 2a) for an emotion with a cognitive constituent, at least one of the cognitive constituents is in part inferable from the inferential base or; 2b) For a cognitive or non-cognitive emotion, a corresponding cognitive prototype rule must be held consciously or subconsciously by the agent that states that the expression of the influencing emotion in the type of moral circumstance at hand is a potential factor that may lead one to make the relevant categorization. Furthermore, this rule must be able to be in part inferable from the inferential base. If i) children can experience moral emotions such as guilt and anger that are typically thought to influence moral judgment in certain cases, ii) developmental psychologists hypothesize that emotions at times influence children's and adolescents' moral judgments, iii) it prima facie experimentally has been shown that emotions can at times influence adult moral decision-making, then an inference may be made that there is a strong possibility that future studies on children will demonstrate that at times they may be influenced by emotions in the appropriate qualified way when making moral judgments.

5. Induction, Concept Combination, & Strong Pluralism

In this chapter, we will tie up the remaining loose ends by examining the prototype, exemplar, theory, and emotion theories of moral concepts in light of the induction and concept combination desiderata. For these desiderata, the specified qualifications from the previous chapter still hold. Recall that the general appropriate qualified way in which a mental representation influence can be said to constitute the concept at hand is when the influence is knowledge that is used by default in moral cognition and is at least in part inferable from the concept in question's inferential base. Also, remember that if the three retraction, inferential base, and informational value tests are not run on participants to examine this specification, a prima facie concept constitution claim may still be made so long as participants do not appear to be experiencing a mental error; a mental error that itself may be caused by various things such as being hit on the head, mistaken reasoning, mental fatigue, being under the influence of drugs, etc. Recall that most concept constitution studies in the concrete and moral concepts literature do not run the above three tests. However, the enormous number of such studies need not therefore be discarded. Rather, such studies can be said to provide initial positive evidence, where most constitution claims in the overall concepts literature are prima facie ones. Next, Fodor's concept combination objections against the prototype view will be addressed. Fodor provides several objections primarily against the prototype theory that prototype concepts do not combine such that they can account for compound thoughts. I will attempt to address Fodor's objections by

relying on my tetrad view of moral concepts. Then we will discuss how my overall theory of moral concepts puts forth a quadruple process theory of moral judgment, where there are four different kinds of cognitive processes each respectively associated with the four different types of moral concept structures. This is contrasted with, for example, Joshua Greene's dual process view in which moral cognition is underwritten by reasoning and emotions-based processes. Finally, by drawing on Weiskopf, we will see why a strong pluralism rather than a hybridism or a weak pluralism applies to the four viable theories of moral concepts, where strong pluralism, hybridism, and weak pluralism were initially introduced in the first chapter.

5.1 Category Induction

Recall from the first chapter that the induction criterion is where theories of concepts apparently must have a structure that allows and can explain how we make inductive inferences about the world. As induction is a higher cognitive competence that is reliant on concepts, an ability to account for this criterion initially appears to be important for any theory. While there are various kinds of induction, what is meant by "category induction" in the concepts literature is the broadening of information from an already known *given* category to a new action or object that falls within that category, or it may also involve such broadening from a previously known given class to a known but different or *target* category. An instance of the former is that if one has knowledge about *right action*, where acts that fall under this class are ones that a person should perform,

and I give you new advice on a moral matter that a particular action is right, such as stopping global warming, one may infer through induction that this act is the act that one ought to perform. An example of the latter is where if I am told that those classified under *humble person* have the property of being soft spoken, then I may induce that those under the target class *meek person* also have this property. Category induction will be the sole type of induction examined in this dissertation, and any use of the term 'induction' will mean only category induction.

To the best of my knowledge, there are no studies on moral concepts specifically attempting to account for the induction desideratum. Even though this is the case, we may still mount enough evidence to infer that the prototype/exemplar and theory views are most likely responsible in the specified way for many instances of induction. By drawing on previously discussed moral categorization studies as well as on induction test results taken from the concrete concepts literature, a prima facie case will be made that the prototype/exemplar and theory views satisfy the induction desideratum for theories of moral concepts.

In the concrete concepts domain, similarity-based accounts have been shown to be responsible in the qualified way for induction in many instances.[1] While most of such induction studies have presupposed that the prototype theory is qualifiedly at work and

---

[1] For similarity-based inductions studies, see Lance Rips, "Inductive judgments about natural categories," *Journal of Verbal Learning and Verbal Behavior* 14: (1975), 665-581. D. Osherson, E. Smith, O. Wilkie, A. Lopez, and E. Shafir, "Category-based induction," *Psychological Review* 97: (1990), 185-200. S. Sloman, "Feature-based induction," *Cognitive Psychology* 25: (1993), 231-280.

there have been very little experiments as of yet explicitly testing for the applicability of the exemplar theory,[2] recall from the third chapter that it is very difficult to differentiate whether one's results show the influence in the appropriate specified way of prototypes or exemplars since one's complete set of exemplars will also contain representations of all the statistically frequent general features that are represented by one's prototypes. Thus, even though psychologists have presupposed the prototype theory in their induction models, this does not necessarily mean that exemplars are not being used in the qualified way. Given the uncertain state of the current induction literature for concrete concepts for prototypes and exemplars, any results showing the qualified influence of some kind of similarity-based account for concrete concepts could be evidence for the use of prototypes or exemplars. To mark this uncertainty, I state that the '*prototype/exemplar*' theory prima facie may be shown to satisfy the induction desideratum.

In order to first make our case, we will examine induction studies in the concrete concepts literature. In Lance Rips' classic induction studies, he used blank predicates.[3] Blank predicates are predicates that subjects will most likely be unfamiliar with, attributes that most people will not have any strong feelings about, or qualities that are made up. Using blank predicates, such as 'sesamoid bones,' allows experimenters to test for induction without any influence of previous knowledge about the attribute. In one test, participants were asked how strong the induction is that if robins have sesamoid

---

[2] Murphy, *The Big Book of Concepts*.
[3] Rips, ibid.

bones, then all birds have sesamoid bones. Another proposed induction was that if penguins have sesamoid bones then all birds have sesamoid bones. Rips found that an induction is considered to be stronger to the extent that the given category is judged to be more similar to the target category. Since the given category of robins is more similar to the target category of birds than penguins are to birds, subjects judged that the induction from robins to birds is stronger than the induction from penguins to birds. Moreover, another factor that influences induction in the specified way is typicality. If the given category is a more typical member of the target category as compared to a different given category, then this will lead to a stronger inductive inference. Since robins are more typical birds than penguins, the induction involving robins is stronger than that which involves penguins. Insofar as the prototype and exemplar theory can account for similarity and typicality effects, this study prima facie suggests that prototypes/exemplars are at work in the specified way in certain cases of category induction.

While demonstrating that similarity plays a role in inductions, Osherson and company have further shown that *coverage* can also influence induction judgments in the appropriate specified way.[4] The notion of coverage is dependent on how well the objects in the premises cover the smallest category that includes all the items in the induction problem; the greater the level of coverage of the smallest inclusive category, the stronger the perceived induction. Take for example:

---

[4] Osherson, ibid.

1) Foxes have sesamoid bones.
Pigs have sesamoid bones
Gorillas have sesamoid bones

2) Foxes have sesamoid bones.
Pigs have sesamoid bones
Wolves have sesamoid bones.
Gorillas have sesamoid bones.

In both 1) and 2), the smallest inclusive category is mammals. 2) has a larger coverage of the mammal category than 1), so this explains why participants generally understand 2) to be a stronger induction than 1).

Regarding the theory-theory, Proffitt and company have demonstrated that at times underlying causal explanatory knowledge may in the specified way be responsible for inductions rather than superficial similarity, typicality, or coverage.[5] For instance, tree experts were told that a blank predicate disease A affects a tree species *x*, while disease B affects a different tree species *y*. The experts then were asked which of the two diseases would most likely affect the other kinds of trees found in their local area. Our experimenters discovered that tree experts were more likely to base their judgments in a qualified way on knowledge of causal explanations rather than superficial similarity, typicality, or coverage. Experts surprisingly would draw conclusions that at times went counter to the conclusions that would be drawn if participants were relying on superficial similarity, typicality, or coverage. Rather, they would base their inductions on theory

[5] J. Proffitt, J. Coley, D. Medin, "Expertise and category-based induction," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: (2000), 811-828. Also see A. Lopez, S. Atran, J. Coley, D Medin, and E. Smith, "The tree of life: Universal and cultural features of folkbiological taxonomies and inductions," *Cognitive Psychology* 32: (1997), 251-295.

knowledge. For example, tree experts would in the specified way use knowledge of the root systems of trees and whether certain trees had root systems that were more likely to spread diseases. This provides prima facie evidence for the theory-theory in induction.

Rehder has found that whether similarity or theory knowledge is qualifiedly used in induction can at times be dependent upon the learning circumstances and the individual.[6] Subjects were told of an imaginary object called a 'Romanian Rogos,' and they were told four properties that seventy-five percent of Romanian Rogos have, such as having an engine. Participants were then told of a new Romanian Rogo that possesses a new property. Half the subjects were told of an underlying explanation to the new property. For instance, the Rogo was said to have the new property of melted wiring, but this was due to the fact that the Rogo has a really hot engine. The other half of the participants were not told this explanatory theory knowledge. The experimenters also varied the typicality of the new Rogo on different vignettes by varying the number of the four common Rogo properties the new Rogo has. Subjects were then asked that in light of the new Rogo with the new property, how likely was it on a scale that all the other Rogos would have the new property. Rehder found that typicality qualifiedly influenced induction judgments made by those who were not told the underlying causal explanation. However, for those subjects who were informed of the theory knowledge, half of them qualifiedly used such knowledge in their inductions. The other half used similarity

---

[6] B. Rehder, "When causality and similarity compete in category-based property induction," *Memory & Cognition* 34: (2006), 3-16.

knowledge in the specified way when making their judgments regardless of the fact that they were told about the causal explanation. This demonstrates that factors such as learning conditions and the subjective tendencies of the particular individual may prima facie determine whether similarity or theory knowledge is qualifiedly used.

The above studies prima facie establish that prototype/exemplars and theory knowledge are at work in the specified way at various times in the higher competence of drawing inductions for concrete concepts. Previous chapters of this dissertation have also established that based on categorization studies there is prima facie empirical evidence that moral concepts can have prototype, exemplar, or theory structures, amongst others. Moreover, humans do in fact make moral inductions all the time. Remember our previous example of one type of category induction that if one has knowledge about *right action,* where acts that fall under this class are ones that a person should perform and I give you advice in a moral matter that a particular action that is new to me, such as stopping global warming, is right, one may induce that this act is the act that one ought to perform. Since we do draw moral inductions, the knowledge stored in our concepts must be doing the work in these cases. As it has been prima facie shown that prototypes/exemplars and theories in varying circumstances qualifiedly do the work for concrete concepts in induction cases and there is prima facie evidence that moral concepts can have prototype, exemplar, or theory structures from categorization studies, we may draw an inference based on theoretical simplicity that at least prototypes/exemplars and theories in various circumstances also do the specified work

for moral concepts in moral inductions. Given the available evidence that such structures are responsible for induction for concrete concepts, there is no reason to complicate the story by supposing that things are different for the moral concepts cases. Notice that this inference from simplicity is based on empirical evidence from moral and concrete concepts. Since such structures do the induction work in the specified way for concrete concepts and there is evidence for such structures in moral concepts, this presents the prima facie case that such structures at least are also qualifiedly responsible for moral inductions.

For example, using blank predicates like 'property x,' one may be presented with the problem of judging on a scale how strong the induction is from 'brave people have property x, so therefore, good people have property x.' Another case may be 'just people have property x, so therefore, good people have property x.' When presented with these problems, a subject may judge the second to be a stronger induction, and the explanation for this may be that the participant was in large part qualifiedly using similarity-based knowledge, such as prototypes. For, just people are more similar to, typical of, and have a greater degree of coverage of good people as compared to brave people for this subject. On the other hand, another participant may draw the same conclusion, but this participant may have qualifiedly used folk consequentialism theory knowledge. When given the vignette, this participant may have reasoned that just people will bring about better consequences for society than persons who are only brave. Due to this inference and the fact that this subject maintains that good people bring about the best consequences for

society, this participant judges that the induction involving just people is stronger than the one involving brave people.

There is the further issue of whether prototypes or exemplars are qualifiedly used in moral inductions. As it is unclear which of these structures are so used in inductions in the concrete concepts domain and I use induction studies for concrete concepts in order to in part draw conclusions for moral concepts and moral inductions, the above ambiguity must also hold in the moral concepts domain as well. Resolution of this ambiguity may be important because induction is a moral concept theory desideratum that any moral concept theory purportedly must attempt to satisfy. If, for instance, prototypes rather than exemplars are really responsible, in the specified way, for moral inductions, then this appears to be problematic for the exemplar theory. Likewise, for the emotion theory, remember from the second chapter that within constraints, I do not make a definitive claim as to what emotions are, and I allow for the possibility that emotions may be non-cognitive, where emotions are not constituted by appraisal judgments related to one's well-being. If emotions are cognitive and constituted by appraisal judgments, then emotion theory moral concepts may be able to handle the induction desideratum since the satisfaction of this desideratum requires that moral concepts decompose into further concepts that, for example, allows for the measurement of such things as similarity and typicality effects. However, if some emotions may be non-cognitive, then it is wholly unclear how such emotion theory moral concepts may be qualifiedly responsible for

moral induction. Since emotion theory concepts may not satisfy the induction desideratum, this potentially may be problematic for the emotion theory as well.

What we have here is the problem that it may be the case that some theories of moral concepts, such as the exemplar and emotion theories, may not be able to handle the induction criterion. For, given the ambiguity in the current data between the prototype and exemplar theories, it may turn out in the future that really prototypes rather than exemplars satisfy this desideratum. Moreover, some emotions may be non-cognitive. As we shall see in the next section, a similar potential problem may also hold for the desideratum of concept combination. The response to this worry is that we should relax the desiderata for theories of moral and concrete concepts. The idea that a single theory of concepts needs to satisfy all the desiderata of a theory of concepts is a by-product of the days when philosophers and psychologists thought that there could only be one viable theory of concepts. For example, the British Empiricists generally believed that all concepts are pictorial images. As discussed in previous chapters, now that there is strong empirical evidence that there may be a multiplicity of viable conceptual structures in the moral and concrete domains that may be qualifiedly used in different circumstances and times, we no longer have to perceive the concept theory desiderata as being inert edifices marking the guidelines for determining what single theory of concepts is correct. Some concept structures may be responsible in the specified way for competences such as induction while different structures qualifiedly may be at work in different higher competences and in different situations. Relaxing the desiderata where a theory of

concepts need only satisfy some rather than all of the desiderata is more appropriate, and, in light of the prima facie empirical evidence where there is a multiplicity of conceptual structures at work in the specified way in higher cognition in different circumstances and times, it brings a coherence to our meta-theory of how we should theorize about theories of moral and concrete concepts. To note, this may begin to sound like an implicit statement of strong pluralism.[7] However, I will contend for strong pluralism at the end of this chapter. Now, even if the emotion theory and, for example, the exemplar theory cannot handle the induction criterion, this does not necessarily mean the demise of such views. Emotion and exemplar knowledge still participate in the specified way in moral categorization and in potentially other competences, while prototypes and theories are qualifiedly responsible for moral inductions, *inter alia*. In these respects, my overall theory of moral concepts still stands. As concepts are mental representations that play a qualified functional role in some of the higher competences, it has already been prima facie shown that emotion and exemplar bodies of knowledge do play such a role in making moral categorization judgments. To contrarily conclude that the emotion and

---

[7] Strong pluralism allows it to be contingently true that if one type of knowledge of *x* is used in a particular competence, then another body of knowledge of *x* may be used in a different competence. However, the fact that different bodies of knowledge may partake in some competences rather than in others may not necessarily lead to strong pluralism, although it may begin to sound like strong pluralism. One reason why it does not necessarily lead to strong pluralism is that even if emotion and exemplar knowledge participate in categorization rather than induction, it could be the case that some other body of knowledge, such as theory knowledge, may partake in all the various higher competences, and theory knowledge is the competently mastered criterion of correctness in all the higher competences. Therefore, hybridism is still a live possibility.

exemplar theories for moral concepts are not viable because they potentially cannot handle the induction desideratum would be to conclude that none of our moral concepts are constituted by such knowledge and therefore, such knowledge does not partake in moral cognition at all; a gross misrepresentation of the empirical facts. Given the relaxation of the desiderata, we may understand Machery's previously given definition of a concept to be altered such that concepts are bodies of knowledge that need only be used by default in some rather than in most of the higher cognitive competences.

5.2 Concept Combination

Recall from the first chapter that concepts purportedly must respect productivity and systematicity.[8] Productivity is where with a finite set of concepts and rules of combination, we can have an infinite amount of new thoughts. For example, if we have the concept AND with the letters of the alphabet, we can have A AND B, A AND B AND C, AA AND B AND BCC, etc. ad infinitum. Systematicity is where if one understands the given concepts and combination rules, one may understand variations on a thought. An example of this is where if one can understand JACK KISSED JILL, then one may also understand the idea JILL KISSED JACK.

Concept combination is the desideratum that attempts to explain how lexical or even compound concepts in many cases combine with each other in order to form even more complex concepts or ideas. It examines how certain mental representations in many

[8] Fodor and Pylyshyn, ibid.

cases play the specified causal role in how we are able to understand at times new compound thoughts. Combination apparently must respect the productivity and systematicity of thought. If the constituent thoughts of a compound concept come from the constituent thoughts of the lexical concepts that make up the compound, then, along with the rules of combination, productivity and systematicity will be respected. Concept combination apparently must also account for emergent features, or supposed new components of a compound concept that represent features that apparently "emerge" from a complex class that are not ordinarily thought to be contained within the constituent classes. As an example of prototype concepts accounting for concept combination, in Smith, et al.'s classic selective modification model for concept combination as well as in some of the other models that can be viewed as advancements upon it in that, for example, these advanced models account for emergent features, prototype components of the constituent concepts qualifiedly combine in certain cases to account for the components of the complex concept.[9] For instance, in the adjective-noun phrase SMOOTH APPLE, the prototype components of SMOOTH and APPLE combine with each other in the specified way to account for the components that make up the compound SMOOTH APPLE. In order to account for concept combination, most lexical concepts

---

[9] E. Smith and D. Osherson, "Conceptual combination with prototype concepts," *Cognitive Science*, 8: (1984), pp. 337-61. J. Hampton, "Inheritance of attributes in natural concept conjunctions," *Memory & Cognition*, 9: (1987), pp. 55-71. D. Medin and E. Shoben, "Context and structure in conceptual combination," *Cognitive Psychology*, 20: (1988), pp. 158-90. G. Murphy, "Comprehending complex concepts," *Cognitive Science*, 12: (1988), pp. 529-62. E. Smith, D. Osherson, L. Rips, and M. Keane, "Combining prototypes: A selective modification model," *Cognitive Science*, 12: (1988), pp. 485-527. J. Prinz, *Furnishing the Mind*.

apparently should have a structure of being constituted by further concepts so that the structures of lexical concepts qualifiedly combine with each other when their respective lexical concepts are joined together to form a compound. This allows us in certain cases to understand new compound thoughts. As an example of concept combination, since SMOOTH is made up of SURFACE and NO ROUGH SPOTS while APPLE is constituted by FRUIT, RED, MEDIUM SIZED, and PICKED FROM TREES, the components of SMOOTH and APPLE causally combine in the specified way in order to account for the compound thought SMOOTH APPLE. To the best of my knowledge, there are no direct studies on concept combination and moral concepts. However, a like argument will be made as with the above induction desideratum in order to prima facie show that at least some of the viable theories of moral concepts can account for this desideratum. The first step in being able to show this is to examine concrete concepts and concept combination.

Hampton found that the typical features of the members of the individual categories that are represented by the constituent concepts in a compound are also the typical features found in the members of the compound category that are represented by the compound concept itself.[10] For instance, the typical feature of *no rough spots* for *smooth* and the typical features *fruit*, *red*, and *picked from trees* for *apple* for me are also the typical features of what I think belong to *smooth apple*. As the exemplar theory can also be qualifiedly responsible for these results of conceptual combination, Hampton's

---

[10] Hampton, ibid.

study provides prima facie evidence that prototypes/exemplars are responsible in the specified way for concept combination.

Smith and Osherson discovered a conjunction effect in participants in concept combination.[11] This is where a compound item such as a red apple is deemed to be more typical of the corresponding compound category than the relevant individual component categories. For example, a red apple is more typical of *red apples* than of *redness* or *apples.* A brown apple is generally categorized as being somewhat typical of *browness* but not typical of *apples*. However, it is judged to be very typical of *brown apples*. The interesting question with the brown apples case is how something not typical of apples can be very typical of a subtype of apples.

To explain this, Smith and company used a schema or prototype structure in their selective modification model. Once again, however, exemplars could be used in the specified way in this model and produce the same results, although for now, we will explain this model in terms of the prototype theory. A schema concept contains representations of dimensions and features for those dimensions. For instance, APPLE may contain the dimensions COLOR, SHAPE, and TEXTURE. These represented dimensions are weighted in terms of importance. In our example, COLOR may carry more weight for me than SHAPE and TEXTURE. However, the weight for the dimension representations may change based on context. For instance, if our discussion solely focuses on the shape of apples and not on its color or texture, where I token my CIRCULAR APPLE compound, then

---

[11] Smith and Osherson, ibid.

the represented dimension of SHAPE now carries more weight than COLOR for the compound concept CIRCULAR APPLE. Under the dimensions, such as COLOR, there will be representations of features, such as RED, GREEN, and BROWN. These feature representations also carry weights, and the weights can change based on the context. For example, RED and GREEN may carry significant weight for me, but BROWN carries little weight for my APPLE concept. However, if our conversation is about brown apples, where I token my BROWN APPLES compound, then the weight for BROWN sharply increases in my compound concept BROWN APPLES.

When we have a conversation about red apples, more weight for RED APPLES will be shifted to the dimension representation COLOR given the context of our discussion on red apples. Also, the feature representation RED will further dramatically increase in its weight for my RED APPLES compound given the focus of our conversation. This explains the conjunction effect in that the weight increase for the relevant dimension and feature for the compound concept partially accounts for why a red apple is considered to be more typical of *red apples* than of *apples*. First, since REDNESS does not contain many or any mental representations related to fruit but RED APPLES do, the similarity score of a red apple to RED APPLES is greater than the score of a red apple to REDNESS. Therefore, a red apple is considered to be more typical of *red apples* than of *redness*. Moreover, since there is the matching of colors between red apples and RED APPLES and there is added weight to COLOR and RED in RED APPLES, where unlike the APPLES concept, all the weight in the COLOR dimension gets placed on RED for RED APPLES due to the combination of RED

with APPLES, a red apple has a higher similarity score for RED APPLES than just for APPLES. This is why a red apple is considered to be more typical of *red apples* than *apples*. This likewise holds for why a brown apple is considered to be a more typical *brown apple* than an *apple*. There is the matching of colors between brown apples and BROWN APPLES. Furthermore, the drastic weight increase for the represented dimension COLOR and the represented feature BROWN in BROWN APPLES when we discuss brown apples creates a higher similarity score of a brown apple to BROWN APPLES than to APPLES.

A virtue of the selective modification model is that it can account for this conjunction effect. Empirical confirmation of its ability to account for this effect provides prima facie evidence for the prototype/exemplar theory in concept combination. However, there are problems with this model when assumed to being a prototype model in that, for instance, it cannot properly account for emergent features.[12] We soon will discuss newer concept combination models that can account for emergent features in certain instances, but broadly speaking, the newer models still have had to implement Smith and company's general idea of advancing weights in order to account for the conjunction effect. In this respect, the selective modification model has longevity.

There are prima facie studies showing the qualified influence of exemplars in concept combination in order to account for emergent features, such as by Hampton and Medin & Shoben.[13] For example, PET FISH for many people has the component TROPICAL

---

[12] For objections to the selective modification model, see: Murphy, ibid. Medin and Shoben, ibid.
[13] Hampton, ibid. Medin and Shoben, ibid.

COLORED. However, based on a prototype view, this component may not appear to be a constituent of PET or FISH when taken individually. The feature may not be in a summary or general representation of features for me for *pet* or *fish*. It somehow appears to mysteriously emerge from the combination of the concepts, but this cannot be the case if concepts combine through strict compositionality. Nevertheless, the exemplar theory can qualifiedly account for the apparent emergent feature. For, one has stored exemplars of pets or of fish that are of tropical colored fish. For instance, I've seen a tropical colored fish in a fish tank at the dentist's office. I have not seen enough tropical colored fish for TROPICAL COLORED to be a part of a relevant summary representation of features, but I still have a stored exemplar of such a fish or of such a pet. The corresponding representation of this particular instance is stored as an exemplar of a particular pet or of a particular fish. Since the exemplar itself is constituted by representations of features such as TROPICAL COLORED, the apparent emergent component of the compound concept does actually come in the qualified way from the relevant constituent concepts of the compound.

Likewise, for WOODEN SPOON, LARGE may appear to be an emergent component not contained within the relevant prototype bodies of knowledge. However, I have seen a limited number of large wooden spoons in kitchens so I have exemplar mental representations of particular wooden things or of spoons that are about particular large wooden spoons I have encountered in my life. As the relevant exemplar is itself constituted by components that represent features such as *large*, the apparent emergent

component of the compound WOODEN SPOON does come in the specified way from the relevant constituent concepts.

There are also prima facie studies on concrete concepts showing the work of theory knowledge in the specified way in concept combination. On Johnson and Keil's theory knowledge model, they found that participants listed many apparent emergent features for certain compounds.[14] For instance, for *arctic bicycle*, subjects listed the feature *spiked tires* even though they did not list this feature for *arctic* or *bicycle*. Upon finding these results, they then ran other studies asking participants to fill in the blanks of, for example:

> Since the ARCTIC _____
> and since BICYCLES_____
> then arctic bicycles have spiked tires.

They found that participants filled in the blanks generally with longer phrases that were not included in the prototype feature-listing task and that contained causal background theory knowledge. For example, one participant said that bicycles "require traction to move." Then the experimenters asked the subjects to rate how likely supposed emergent features would belong to a compound category before and after the removal or denial of the relevant theory knowledge. For instance, subjects were asked to rate how likely arctic bicycles would have spiked tires. Then they would be informed that arctic bicycles are jet propelled. Subsequently, they would be re-asked to rate how likely arctic bicycles

---

[14] C. Johnson and F. Keil, "Explanatory understanding and conceptual combination," in *Explanation and Cognition*, ed. By F. Keil and R. Wilson, (Cambridge, MA: MIT Press), 2000, 328-359.

would have spiked tires. They found that participants substantially lowered the likelyhood that the compound category would still contain the supposed emergent feature. This study provides prima facie evidence that at times, theory knowledge is qualifiedly at work in concept combination for concrete concepts.

The above studies establish the prima facie claim that depending on the circumstances, exemplars and theory knowledge are at work in the specified way in concept combination for concrete concepts and perhaps prototypes are as well.[15] Previous chapters of this dissertation have also prima facie established that based on categorization studies, there is empirical evidence that moral concepts can have prototype, exemplar, or theory structures, amongst others. Moreover, humans do in fact combine moral concepts all the time. There are compound moral concepts in which two constituent concepts are moral or thick moral concepts, such as: MORAL SAINT, CHIVALROUS GENTLEMAN, VIRTUOUS LADY, DEVIOUS CROOK, HONOR AMONGST

---

[15] Murphy argues that the empirical fact that compounds like TYPEWRITER TABLE and CORPORATE STATIONARY combine based on a similarity view shows the working of prototypes rather than exemplars because such compounds are not the intersection of thoughts about, for example, representations of a particular typewriter and a particular table or representations of a particular corporation and a particular stationary. However, I believe that there is still the possibility that the exemplar theory can account for combination in the above cases. For TYPEWRITER TABLE, I have exemplars stored in my TABLE concept that refer to particular typewriter tables. These exemplars will be contextually triggered by the combination of TABLE with TYPEWRITER. The exemplars each have components that are responsible for the component thoughts I have concerning TYPEWRITER TABLE. An analogous story may be told for CORPORATE STATIONARY as well. This demonstrates that until there are definitive discriminating studies between the prototype and exemplar theories for conceptual combination, positive prima facie similarity-based findings in combination studies suggest the qualified use of prototypes/exemplars. Gregory Murphy, *The Big Book of Concepts*.

THIEVES, KIND BENEFACTOR, and MALICIOUS CRIME. Moreover, there are compounds that contain only one moral concept, such as: SENSIBLE KNAVE, BRAVE WOMAN, JUST STATEMENT, IMMORAL CHILD, and GENEROUS MAN. Since we frequently do combine moral concepts, the knowledge stored in our concepts must be doing the specified work in these cases for how we understand complex thoughts. As it has been prima facie shown that exemplars, theories, and perhaps prototypes in varying circumstances do the qualified work for concrete concepts in concept combination cases and there is prima facie empirical evidence that moral concepts can have prototype, exemplar, or theory structures from categorization studies, we may draw an inference based on theoretical simplicity that at least exemplars, theories, and perhaps prototypes prima facie also do the specified work for moral concepts in moral concept combination. Given the available evidence that such structures may be at work in concept combination for concrete concepts, there is no reason to complicate the story by supposing that things are different in the moral concepts case. Notice that this inference from simplicity is based on empirical evidence from moral and concrete concepts. Since such structures prima facie do or (for prototypes) may do the specified combination work for concrete concepts and there is evidence for such structures at work in moral categorization, this presents the case that such structures at least prima facie are or may be also responsible in the specified way for moral concept combination.

For example, for MORAL SAINT, the typical features I may list are *kind*, *honest*, *perfect*, and *beneficent*. However, the corresponding prototype components qualifiedly

are combined from either of the components of my two relevant constituent concepts. For VIRTUOUS LADY, VIRTUOUS contains exemplars such as MY BROTHER HELPING THE POOR YESTERDAY, MY SISTER BEING KIND TO THE STRANGER LAST WINTER, and MOTHER THERESA. LADY is constituted by exemplars of females who are considered to be lady-like, like MY MOTHER and QUEEN ELIZABETH. The constituent exemplars are qualifiedly responsible for why when I token my VIRTUOUS LADY concept, I have the thoughts MOTHER THERESA and MY MOTHER. For JUST STATEMENT, JUST for me is constituted by utilitarian theory knowledge. This knowledge is causally responsible in the specified way for the thought MAXIMIZE THE GREATEST HAPPINESS FOR THE GREATEST NUMBER when understanding the complex idea JUST STATEMENT and when bringing forth this compound into working memory.

Concerning the emotion theory and concept combination, there may be a problem with some concepts being constituted by emotions in that it may be the case that some emotions at times may not be made up of appraisal judgments. Thus, as combination is seemingly dependent upon concepts decomposing into further concepts, where cognitive emotions may be candidates for concept combination, there is the issue of whether concepts constituted by non-cognitive emotions may actually combine and fulfill this concept combination desideratum for theories of moral concepts. This is especially a very serious potential problem for the emotion theory since if emotions are non-cognitive, then it is difficult to see how emotion theory concepts can have the satisfaction conditions or the formal properties that allow for concept combination. The response to

this worry in relation to emotions that may be non-cognitive is that once again, a theory of concepts need not account for all the desiderata. We have seen how there are several different viable conceptual structures generally at work in the specified way in the human mind in the relevant higher competences in moral cognition; various kinds of knowledge whose recruitment may be context sensitive. However, an interesting facet of having a multitude of different structures of a class is that at many times in cognition, several different kinds of knowledge of a particular category may qualifiedly work together in cognition. For example, Frank Keil and company show that prototype and theory types of knowledge frequently work together in the specified way in the classification of particular natural kinds.[16] Given that different kinds of knowledge frequently can so work together, for emotions that may be non-cognitive, the emotion body of knowledge and another type of knowledge one has of *kindness* may qualifiedly work together in tandem in cognition. To note, while a hybridist and strong/weak pluralist may potentially make this claim, I will discuss these views at the end of this chapter. Since it has been prima facie inferred that at least the exemplar and theory structures work in the specified way in concept combination, the concept combination objection against the emotion theory may at least be initially handled in that the potentially non-cognitive emotion-based knowledge and some other kind of knowledge such as exemplars of the same particular category *x* may work together in the specified way in a single act of reasoning,

---

[16] F. Keil, W., Smith, D., Simons, D., Levin, "Two dogmas of conceptual empiricism: implications for hybrid models of the structure of knowledge," *Cognition*, 65: (1988), pp. 103-135.

where in such cases, the exemplars can account for the concept combination desideratum and the emotion-based knowledge can simultaneously still play a qualified role in, for example, categorization or planning. The potential non-cognitive emotion theory knowledge need not take part in the specified way in concept combination and possibly not in other higher competences as well. Furthermore, if it is later shown that prototypes do not play the qualified role in combination, then this same move may be made for the prototype theory. Prototypes and, for instance, exemplars may work in the specified way in tandem in cognition, where exemplars account for combination, and prototypes are responsible for categorization in certain cases.

5.3 Fodor's Concept Combination Objections Against the Prototype Theory

Jerry Fodor argues that lexical concepts rarely decompose into components that represent non-defining features, such as prototypes.[17] His justification that lexical concepts usually do not decompose into representations of non-defining features is a series of objections primarily against the prototype theory that it cannot handle concept combination.

The first objection claims that the prototype theory cannot account for special phrasal concepts such as: AMERICAN CITIES SITUATED ON THE EAST COAST JUST A LITTLE SOUTH OF TENNESSEE, GRANDMOTHERS MOST OF WHOSE GRANDCHILDREN ARE

---

[17] Jerry Fodor, "The Current Status of the Innateness Controversy," *Representations*, (Cambridge, MA: The MIT Press), 1981. Fodor, *Concepts*.

DENTISTS, and NOT A CAT. Fodor states that the component concepts of the compounds supposedly do have prototype structure, but the compound concepts themselves do not. Prototypes do not combine. Constructing analogous Fodorian examples that in part use moral concepts, we may have: SIN CITIES SITUATED ON THE EAST COAST JUST A LITTLE SOUTH OF TENNESSEE, HIGHLY MORAL GRANDMOTHER MOST OF WHOSE GRANDCHILDREN ARE DENTISTS HELPING THIRD WORLD CHILDREN, and NOT A CRIME.

Whether or not moral and concrete prototype concepts can combine with each other is ultimately an empirical question. Although there is no experimental data on the Fodorian moral compounds I have just constructed above as well as on other Fodorian moral compounds we will discuss later, we have seen evidence in which a prima facie inference may be made that prototypes may combine in the specified way in particular circumstances. Furthermore, there is support for the prima facie inference that exemplars and theories do combine in the qualified way in certain cases. Since there are no studies on my constructed Fodorian moral compounds, we may attempt to address Fodor in a more speculative manner in this chapter section, although it is not completely speculative since I prima facie have already shown based on categorization studies that moral concepts can have prototype, exemplar, theory, or emotion structures and that the aforementioned four bodies of knowledge prima facie do or may potentially partake in concept combination in the specified way depending on the situation. Furthermore, the somewhat speculative manner is warranted in that Fodor just assumes that his special phrasal concepts do not have prototypes without any empirical support.

The first line of response is motivated by Prinz in his work on concrete concepts.[18] If the phrasal concepts are candidates for typicality effects, where some members of the corresponding category are taken to be more typical than others, then this provides suggestive evidence that the phrasal concepts do have prototypes or, in other words, are constituted by a summary or general representation of features. To note, exemplars can also account for typicality effects. However, either way, whether prototypes or exemplars account for typicality effects, since my overall theory of moral concepts includes both theories and I have already established that different bodies of knowledge may be at work in the specified way in different circumstances, the perceived existence of typicality effects still supports my overall theory of moral concepts. My overall theory of moral concepts can handle Fodor's objections if there are perceived typicality effects since my overall view includes both prototypes and exemplars. Since my overall view includes the prototype and exemplar theories, where either one of these bodies of knowledge qualifiedly may be at work in cognition in particular contexts, it does not matter whether it is prototypes or exemplars that account for the perceived typicality effects. If, for example, it is prototypes rather than exemplars that qualifiedly account for the perceived typicality effects, then my overall theory of moral concepts allows for prototypes to causally be at work in the specified way in the relevant special cases for such Fodorian phrasal compounds. This same story holds if it is determined

---

[18] Prinz, *Furnishing the Mind*.

that exemplars rather than prototypes qualifiedly account for the perceived typicality effects or if it is found that both so account for the perceived typicality effects.

For SIN CITIES SITUATED ON THE EAST COAST JUST A LITTLE SOUTH OF TENNESSEE, it is plausible that there can be a typicality scale. For example, CHARLESTON, SOUTH CAROLINA may be low on the typicality scale since I perceive it to generally be a sin city on the east coast but it has some conservative and religious elements. MYRTLE BEACH, SOUTH CAROLINA is high on the scale since it is a well-known beach party destination. What accounts for the typicality scale for SIN CITIES SITUATED ON THE EAST COAST JUST A LITTLE SOUTH OF TENNESSEE may be prototype components such as HAS MANY BARS, HAS GAMBLING, HAS BEACH PARTIES, WARM, etc. The perceived existence of typicality effects for this phrasal concept suggests that this compound is constituted by prototypes/exemplars.

For HIGHLY MORAL GRANDMOTHER MOST OF WHOSE GRANDCHILDREN ARE DENTISTS HELPING THIRD WORLD CHILDREN, a highly moral grandmother who has ten grandchildren with nine of them being dentists who help third world children is more typical than a highly moral grandmother who has ten grandchildren with six of them being dentists who help third world children. This typicality scale may be accounted for by prototypes for the compound concept such as GENEROUS, HAS CHILDREN WHO HAVE CHILDREN, MORAL EDUCATOR, KIND, HONEST, etc. For NOT A CRIME, there are many acts that are not crimes, such as: cooking breakfast, helping the poor, walking the dog,

and driving to work. However, as crimes are usually perceived as being immoral acts,[19] I may have a typicality scale where the most typical acts that are not crimes are the most moral ones. For instance, helping the poor and keeping an important promise are highly typical of not being a crime. Cooking breakfast and driving to work are only moderately typical as they are not necessarily moral acts, while telling white lies – kinds of lies that I do not hold to be crimes – is least typical. The typicality scale for NOT A CRIME may be accounted for by prototypes such as VIRTUOUS ACT, JUST ACT, KIND ACT, HELPFUL, etc. The perceived existence of a typicality scale in the above examples counters Fodor's objection and suggests that my overall theory of moral concepts can handle his contention.

Furthermore, we may add to Prinz's work on concrete concept combination in that theory knowledge at times can also play a specified role in concept combination for the Fodorian compounds related to Fodor's first objection. My overall theory of moral concepts allows for different bodies of knowledge to qualifiedly be at work in different circumstance, and such bodies of knowledge may also work conjointly with other kinds of knowledge. This can also be a potential response to Fodor in some cases. For instance, for NOT A CRIME, I perceive crimes to be immoral and I am a folk consequentialist. The typicality effects I have for this compound are directly accounted for in the specified way in cognition by my ethical theory knowledge that one must

---

[19] Hampton, "An Investigation of the Nature of Abstract Concepts."

perform those acts that bring about the best consequences for society. Acts that bring about better consequences are deemed to be more typical acts that are not crimes.

Another addition to Prinz is that as mentioned in the second chapter, the prototype theory predicts that there will be ambiguous cases of membership, whereas the classical view does not. Since the summary representation of features are not taken to refer to necessary and sufficient conditions, we should expect subjects to be unsure at times about whether some items should be classified in one particular category or another. If it is expected that there are ambiguous cases of membership, then this provides some further support that the phrasal concepts in question may have prototype structure since prototype theory predicts and explains ambiguous cases of membership.[20] With our particular phrasal concepts, we do see the possibility and likelihood of ambiguous cases of membership, which provides a counter to Fodor. For SIN CITIES SITUATED ON THE EAST COAST JUST A LITTLE SOUTH OF TENNESSEE, I may not be sure whether Wilmington, North Carolina falls within the corresponding category. For, it is a reasonably sized beach city and it has many beach parties which could mean that it is a sin city, but at the same time, from my personal experience, I know there are significant charitable activities occurring in the city. Therefore, I am not sure if it falls in the relevant class or not. Also, recall that in the third chapter it was inferred that moral concepts as a class generally do have prototype structure and do produce concomitant

---

[20] Exemplars also may be able to account for ambiguous cases of membership. However, once again, it does not matter here whether it is prototypes or exemplars that account for the phenomenon since my overall theory of moral concepts includes both views.

typicality effects that provide evidence against the classical view. Thus, the above ambiguity for SIN CITIES SITUATED ON THE EAST COAST JUST A LITTLE SOUTH OF TENNESSEE coupled with the fact that SIN has prototype structure that allows for ambiguous cases of membership, puts forth a case that the compound concept at hand does have prototype structure. For HIGHLY MORAL GRANDMOTHER MOST OF WHOSE GRANDCHILDREN ARE DENTISTS HELPING THIRD WORLD CHILDREN, I may know a grandmother, most of whose grandchildren are the relevant dentists, but I am not sure if she is a highly moral person or not. She has some noble qualities but also some malicious one's as well. Therefore, I am not sure whether or not she falls in the corresponding compound category. Also, as further support that this ambiguity may be the case, recall from the third chapter that we examined a study that demonstrated that HIGHLY MORAL PERSON does have prototype structure and does produce typicality effects. For NOT A CRIME, many people appear to be uncertain about whether things like abortion are not a crime regardless of what the law in their state may say. Also, recall from the third chapter that we examined an experiment that demonstrated that CRIME does have prototype structure and does produce typicality effects such that there may be ambiguous cases of membership for this concept. While ambiguous cases of membership do reinforce the plausibility of a prototype analysis as opposed to a classical definitional analysis, insofar as the prototype theory predicts and explains the existence of ambiguous cases of membership, the strong possible existence of ambiguous cases of membership

provides a counter to Fodor's first concept combination objection that certain compound concepts do not have prototype structure.

Fodor's second objection is that there are the previously discussed emergent components to certain compound concepts such as PET FISH.[21] Another example Fodor and Ernest Lepore like to use is BROWN COW.[22] A brown cow has the property of being dangerous or charging at red capes. However, these prototypical properties are not thought by people to be properties of *brown* or *cow* when considered individually. They are apparent emergent features, but how can the existence of such features be the case if the prototypes of lexical concepts supposedly qualifiedly combine to account for the represented features of compound concepts? For moral concepts, the compound SENSIBLE KNAVE has the component prototype feature KNOWING WHEN ONE WILL NOT GET CAUGHT. However, this component feature does not appear to be contained within SENSIBLE or within KNAVE when considered individually. Thus, it appears that there is a problem of concept combination in certain cases for the prototype theory in the domain of moral concepts as well.

The first response that is available for my overall theory of moral concepts is that we may have stored exemplars that may qualifiedly account for the supposed emergent components. We have already discussed how Hampton and Medin & Shoben have shown that exemplars can account for apparent emergent components in certain instances

---

[21] Fodor, *Concepts*, 102.
[22] Fodor, ibid. Jerry Fodor and Ernest Lepore, *Holism: A Shopper's Guide*, (Oxford: Blackwell), 1992.

such as in PET FISH. This may also work for our particular compound moral concept SENSIBLE KNAVE. I have met a few knaves in my life or seen actors play them in movies, some of whom have had good foresight into knowing whether or not they would get caught. However, I have not had enough experiences with such knaves such that the apparent emergent feature *knowing when one will not get caught* is part of the general features I think belong to *knaves*. Nevertheless, such individuals, when understood as exemplar mental representations of particular knaves, can qualifiedly account for the apparent emergent component of the respective compound in that such exemplars are themselves constituted by concepts such as KNOWING WHEN ONE WILL NOT GET CAUGHT. This may potentially handle the emergent features objection in certain cases.

However, Fodor and Lepore anticipate this exemplar theory move and contend that there are still compounds in which one will not have an exemplar.[23] Even if there are exemplars that can handle the pet fish problem, most people do not have the proper experiences to have relevant exemplars that can in some way handle compounds like PET FISH LIVING IN ARMENIA THAT HAVE RECENTLY SWALLOWED THEIR OWNERS. This particular compound has supposed emergent components such as LARGE and VORACIOUS. Likewise, Fodor and Lepore may claim that the exemplar theory cannot account for apparent emergent components contained in certain compounds that contain moral concepts such as SENSIBLE KNAVE LIVING IN ARMENIA THAT HAS RECENTLY

---

[23] Jerry Fodor and Ernest Lepore, "The Red Herring and the Pet Fish: Why Concepts Still Can't be Prototypes," *Cognition* 58: 253-270.

SWALLOWED A NEW YORK STRIP STEAK WHOLE. This compound has the supposed emergent component of LARGE and VORACIOUS that does not appear to be contained within the lexical concepts that make up the compound and where we do not have relevant exemplars of this particular kind of represented individual.

In response to this, we can bring in the theory-theory model of combination and the appropriate use of theory background knowledge. For example, in order to address the BROWN COW problem, Ned Block claims that we may store the background conditional knowledge in COW that DANGEROUS IF BROWN.[24] This kind of knowledge is consistent with the theory-theory in that it is explanatory knowledge about cows that if a cow is brown, then the reason why it is dangerous is that it is most likely an aggressive male. In this fashion, Block claims that we can appropriately address the apparent emergent qualities in certain cases. Furthermore, Prinz also argues that in certain cases background explanatory theory knowledge can play the specified role in conceptual combination for concrete concepts.[25] For instance, he argues that it is background explanatory knowledge of the world and what it takes for something to swallow a grown human being that qualifiedly accounts for the supposed emergent components to PET FISH LIVING IN ARMENIA THAT HAVE RECENTLY SWALLOWED THEIR OWNERS. Here, Prinz posits his Retrieval, Composition, and Analysis (RCA) model of concept combination for concrete concepts in which the first stage of combination involves the *retrieval* of

---

[24] Ned Block, "Holism, Hyper-analyticity and Hyper-compositionality," *Mind & Language* 8, 1-26.
[25] Prinz, ibid.

relevant exemplars.[26] If this stage is unsuccessful at qualifiedly accounting for apparent emergent features, then in the second *composition* stage, we primarily causally combine in the specified way the prototype bodies of knowledge for the relevant lexical concepts. In the final *analysis* stage, we bring in background explanatory theory knowledge to analyze the compound in the specified way and resolve the apparent conflicts and gaps in the composition stage.

Likewise, for SENSIBLE KNAVE LIVING IN ARMENIA THAT HAS RECENTLY SWALLOWED A NEW YORK STRIP STEAK WHOLE, although the exemplar theory may not work for this instance, KNAVE is a thick concept that has a descriptive and a normative component. A descriptive prototypical feature that knaves have is that of being human. We have background explanatory theory knowledge of human beings that if one can swallow things like a steak whole, this individual must be large and have a large mouth. Moreover, this person must have also been voracious. It is the inclusion of background explanatory theory knowledge that can qualifiedly account for concept combination and supposed emergent features in certain cases.[27] As my overall theory of concepts allows

---

[26] Ibid.

[27] Concerning the use of non-default knowledge in order to account for complex thoughts, as we have seen, regarding the combination desideratum, some theories of concepts need not even account for this desideratum at all. Furthermore, for those theories of concepts that posit knowledge that does partake in combination, they need to only show that they can account for compound thoughts in some cases and not all the time. Leaving the concept combination desideratum loose, where some concept theories only need to show that they can account for compound thoughts in some but not all cases, opens the door and allows for combination to not be invariably compositional, where non-concept-constitutive knowledge also can play a role in combination. It ought to be flagged that this really does mean giving up on compositionality as

for the four various bodies of knowledge to work in the specified way in cognition individually or conjointly on different occasions and in different circumstances, my view has the resources to account for apparent emergent qualities even in circumstances where it appears that certain bodies of knowledge, such as prototypes, relevantly fail.

5.4 The Quadruple Process Theory of Moral Judgment

Joshua Greene posits his well-known dual process theory of moral judgment, where moral judgments can be influenced by reasoning and emotion-based processes.[28] Similarly, Jonathan Haidt's social intuitionist model is also a dual process view, where affectively-valenced intuitions in many cases influence moral judgment, but Haidt also acknowledges that sometimes reasoning can influence judgment.[29] Shaun Nichols also puts forth a dual process theory he calls the sentimental rules view. He believes we have what he calls a 'Normative theory' or a set of rules that prohibit certain behaviors. We also have an emotional system that is relevant to moral judgment. To note, his use of the term 'Normative theory' is not to be confused with the theory-theory of moral concepts in

---

Fodor et al. have understood it. Not all of our compound thoughts come from the combining of concepts.

[28] Joshua Greene, "The Secret Joke of Kant's Soul," *Moral Psychology Volume 3*, ed. by Walter Sinnott-Armstrong, (Cambridge, MA: The MIT Press, 2008). "Dual-process Morality and the Personal/Impersonal Distinction: A Reply to McGuire, Langdon, Coltheart, and Mackenzie," *Journal of Experimental Social Psychology*, 2009.

[29] Jonathan Haidt, "The emotional dog and its rational tail: The social intuitionist approach to moral judgment," *Psychological Review* 108: (2001), 814-834.

which ethical theory knowledge constitutes concepts. Although both of these aspects for Nichols of 'Normative theory' and of an emotional system are dissociable, sentimental rules involve both of these mechanisms, where 'Normative theories' become "affect-backed" when adjoined with the emotional system. Nichols maintains that in most cases the dual processes of 'Normative theory' and emotions causally influence judgments.

While the above theorists fail to draw the link with functionalism in the philosophy of mind in order to see that in certain specified cases, causal influence studies can lead to constitution claims on concepts, the overall view of moral concepts I have propounded will beneficially add to the work put forth in providing the dual process views by giving a more fine-grained account of the reasoning processes involved in moral cognition. This will result in a quadruple process theory. I have shown that there are four different kinds of bodies of knowledge that potentially may constitute a moral concept and that are at work in some of the higher cognitive competences: the prototype, exemplar, theory, and emotion structures. Concepts psychologically are not only in significant part responsible for higher acts of cognition such as decision-making, induction, deduction, categorization, planning, and analogical reasoning, but the knowledge stored in concepts are in part significantly responsible for the kinds of cognitive processes that underwrite the performance of the higher cognitive competences. Here, cognitive processes for the higher competences are a series of operations that access knowledge in long-term memory in order to perform the functions of higher acts of cognition.

As fully discussed in previous chapters, prototypes are bodies of knowledge of a summary representation of things like rules and virtues generally found in moral acts and persons, exemplars are bodies of knowledge of particular acts and persons, theories are bodies of knowledge of ethical theories, and emotion-based knowledge at least in part contains different mental states from the above views. Furthermore, they rely on different cognitive processing, where different cognitive processes may be used individually or conjointly in divergent situations in moral cognition. Recall that the use of prototypes is underwritten by a linear similarity-based model, but exemplars are grounded on a non-linear similarity-based computation. The use of theory knowledge is underwritten by non-similarity-based algorithms, while emotion knowledge is founded on emotion-based processing.

The overall theory of moral concepts I have provided demonstrates that since there are four viable structures of moral concepts that may be used individually or conjointly in disparate acts of cognition, there are four distinct kinds of viable cognitive processes that underwrite moral judgments and moral cognition. While Greene, Haidt, and Nichols are generally on the right track with their dual process models in that they generally say that reasoning and emotions-based processes are at work in moral judgments, we can see that their theories are not fine-grained enough to account for the four discussed moral concept structures with their four distinct kinds of processing that lie at the very heart of moral cognition. A more accurate description of the cognitive

processes, in terms of scientific explanation and prediction, that underlie moral judgments is a quadruple rather than a dual process theory.

The above theorists in question may object that their notion of a 'reasoning process' is a more abstracted category that is at a higher level of generality in which the prototype, exemplar, and theory views are nested. This is just like how apes, dogs, and cats are nested under the more general mammal category. Prototypes, exemplars, and theories are a subset of some abstracted reasoning-based category. Therefore, it is still fine to say that moral cognition is underwritten by a dual process view.

However, as we are discussing cognitive processing for moral cognition and it is concepts that are in significant part responsible for the kinds of cognitive processing that underwrite moral cognition, the more appropriate level of generality for talking about and distinguishing cognitive processing in terms of scientific explanation and prediction is at the more fundamental level of the four moral concept structures. We have seen throughout this dissertation that the four different bodies of knowledge with their four different kinds of cognitive processing can lead to different outcomes in categorization, induction, and concept combination. Even though, for example, prototypes and exemplars are reasoning-based kinds of knowledge, they can lead to different outcomes in, for instance, concept combination. They explain cognition differently and at times provide disparate predictions that are born out. Thus, a quadruple process theory, in accounting for finer distinctions, has a more accurate explanatory power and predictive

success as compared to a general dual process theory that does not differentiate between the disparate kinds of reasoning-based processes.

Relatedly, mathematical psychologists who work on cognitive processing and constructing mathematical cognitive models discuss their subject matter at a more fine-grained level than with the above theorists' rather general description of a reasoning and an emotion-based process. For instance, whether a similarity model is mathematically linear or non-linear is an important difference that can lead to disparate acts in cognition. This is widely understood by mathematical psychologists. Discussing cognitive processing at the supposed higher level of abstraction does not adequately account for explanation and prediction in cognition in the eyes of mathematical psychologists, and it also is not at the same level of accuracy as used by those working on cognitive processing. It is somewhat analogous to a lion and leopard biologist in sub-Saharan Africa always referring to a lion as a 'cat' to her colleagues even though the area also contains many leopards, which are also cats. This leaves her colleagues confused and unclear about her statements. Talk at such a level of generality lacks the requisite degree of accuracy given the specialized biological context. Likewise, the more accurate description of moral cognitive processing in terms of explanation and prediction is the new quadruple process theory entailed by my overall theory of moral concepts. If one is to provide a theory of moral cognitive processing, then one should do so at the appropriate level in terms of explanatory power and predictive success. My quadruple process theory is fine-grained enough to account for important distinctions in the

cognitive processing literature, and for the above reasons, I suggest that it supplant the dual process theories in moral psychology concerning the nature of moral judgment.

5.5 Strong Moral Concept Pluralism

In this final section, we will contend for strong moral concept pluralism rather than hybridism or weak pluralism. Recall from the first chapter that strong pluralism states that we do not have one universal structure for all moral concepts that mentally represent the various moral categories. Rather, for all moral categories, one may have several different concepts of each moral category, where each concept contains a different structure from the others. For instance, I may have four concepts of VIRTUOUS that each contains the different bodies of knowledge of prototypes, exemplars, theories, and emotions, respectively. Hybridism likewise states that there is not one universal structure for all moral concepts. However, hybridism differs from strong pluralism in that for all moral categories, all the different bodies of knowledge of a single moral category are parts of the same concept. Thus, in the example, if VIRTUOUS is a hybrid, then there is only one concept VIRTUOUS that contains four parts that correspond to the four viable structures. Remember that hybridists maintain that concepts are super-structured if and only if 1) when one of the parts is used in a higher cognitive competence, then it is necessarily the case that the other parts of the concept may be used in other cognitive competences; 2) there is a competently mastered criterion of correctness where the parts of a given concept do not produce inconsistent outcomes such

as inconsistent categorization judgments. Strong pluralism denies 1) and/or 2). Recall that there is also an intermediate view or weak pluralism in which one may have several concepts of a single moral category in a particular moral concept subdomain, such as for thin moral concepts like GOOD, BAD, and RIGHT. However, one may have super-structured concepts in another subdomain such as for thick concepts like LIE, CRIME, and PUSILLANIMITY.

One method that is used by Weiskopf to argue for strong pluralism over hybridism and resultantly, the intermediate view is to attack the organization and super-structure of hybrid concepts.[30] Recall in the third chapter that we have discussed many studies prima facie demonstrating that people generally do have prototype knowledge of moral categories. Moreover, in this chapter, the prototype-exemplar chain argument was presented contending that if people have moral prototypes, then they must also have moral exemplars. Also, from our discussion of the theory-theory in the fourth chapter, a prima facie argument for the viability of a folk consequentialism theory knowledge was put forth, where supporting cross-cultural evidence in part suggested that most people have such knowledge about certain moral categories; knowledge that provides the underlying explanation for and systematization of the relevant prototype and exemplar knowledge.

---

[30] To note, an argument against hybridism and weak pluralism is also an argument against multi-hybridism and multi-weak pluralism. Daniel Weiskopf, "The Plurality of Concepts," *Synthese* 169: (2009), 145-173.

Furthermore, there is also a prototype-theory-theory chain argument that is in play for moral concepts. Nelson Goodman points out that to say that two things are similar is to roughly say that they share a certain number of features in common.[31] However, the problem with this is that any two objects can have many properties in common. For instance, two apparently dissimilar things like a hammer and a laptop computer both weigh less than ten million pounds, they did not exist ten billion years ago, they did not exist nine billion years ago, they take up space, they both have functions, they both can be dropped, they can fit in my trunk, they can fit in my house, they can fit in a garbage bag, etc. Due to this fact, Goodman points out that to say that two things are similar requires a theoretical frame of reference in order to delimit and specify what kinds of properties are relevant for determining similarity in a particular instance. Likewise, as suggested by Murphy and Medin, some of our concepts may contain similarity-based knowledge, but there must be at least implicit theoretical constraints on what kinds of similarity-based knowledge gets stored in our relevant concepts such that certain of our mental representations can be said to be similar to objects and things in the world.[32] There must be some kind of mentally represented frame of reference for why constituents like HAS KEYS, HAS A SCREEN, and HAS A HARD DRIVE are similarity-based knowledge I associate with the category *laptop* rather than WEIGHS LESS THAN TEN MILLION POUNDS and DID NOT EXIST TEN BILLION YEARS AGO. Murphy and Medin claim that it is theory

---

[31] Nelson Goodman, "Seven Strictures on Similarity," in *Problems and Projects*, (Indianapolis: Bobbs-Merrill, 1972).
[32] Murphy and Medin, ibid.

knowledge, such as the knowledge of what a computer's function is, that provides the requisite theoretical constraints. Likewise, for moral concepts, why do I have similarity-based knowledge such as HONEST, KIND, and JUST that I associate with the class *good person* rather than A QUALITY HELD BY SOME PEOPLE IN IRELAND, IS NOT EXEMPLIFIED ON SATURN, and IS NOT EXEMPLIFIED ON VENUS? The answer is that there are theoretical constraints that reign in what kinds of similarity-based knowledge get stored in our relevant concepts. More precisely, it is theory knowledge, such as of virtue ethical theory, which provides the reference point for what kinds of similarity-based knowledge is relevant and irrelevant. Virtue ethical theory generally claims that good individuals have and exemplify the virtues; virtues that lead to a flourishing life. Furthermore, we may look to certain highly moral individuals to see what acts should be performed and which virtues should be developed. It is knowledge of this ethical theory that for me may constrain my similarity-based knowledge for GOOD PERSON such that my GOOD PERSON concept is in part constituted by the prototypes HONEST, KIND, and JUST rather than by A QUALITY HELD BY SOME PEOPLE IN IRELAND, IS NOT EXEMPLIFIED ON SATURN, and IS NOT EXEMPLIFIED ON VENUS. For, I believe that being honest, kind, and just will lead one to live a flourishing life. Therefore, we have on our hands a prototype-theory-theory chain argument that if one has prototypes of a moral category, then it follows that one has theories of the moral category as well. The previously detailed evidence for the prototype theory in a variety of moral concepts also provides evidence for the existence of theory knowledge of such corresponding moral categories. This chain argument for

the viability of the theory-theory of moral concepts is consistent with PAC in that it is in part reliant on prototype experimental studies on abstract moral concepts.

In a word, the evidence prima facie demonstrates that most people have at least prototype, exemplar, and theory knowledge of a variety of moral categories such as *moral*. To note, recall that the current evidence for the emotion theory of moral concepts only demonstrates that a limited number of moral concepts prima facie may have emotion structure. Thus, for example, we may understand moral concept hybridism as generally claiming that there is a single concept MORAL that may be made up of at least three and perhaps four different types of knowledge, where emotions would be the fourth. As a result, when the concept is called into working memory, all the various bodies of knowledge are brought into working memory as well, where they are at least made accessible, if not accessed. On the other hand, strong pluralism posits that one may have at least three and perhaps four different concepts of, for instance, the category *moral* that correspond to the relevant viable theories of moral concepts for the concept in question. As a result, all of the bodies of knowledge need not necessarily be brought into working memory, but only the concept(s) that is relevant to or will be used in the cognitive competence at hand will be retrieved from long term memory.

However, Weiskopf points out that one issue about working memory is that it is generally thought to have a limited capacity and cannot hold a large quantity of

information, which puts size constraints on the magnitude of our concepts.[33] Remember that concepts are thought to contain default bodies of knowledge rather than all the possible facts we may know about x. This allows for a significant reduction in the systematic selection of relevant knowledge when we generally reason about x. It appears that at face value, super-structured hybrid concepts are too large to serve as single units of thought even though each part of the concept contains default knowledge. For, a hybrid moral concept brings in all of the different bodies of knowledge into working memory. Given the limited space capacity of working memory, strong pluralism at face value is the more conducive organization of the mind since all the bodies of knowledge need not simultaneously be brought into working memory. Here, I use the term 'at face value' because there does not appear to be much experimental work specifically on concept constitution in relation to the capacity of working memory, so there does not appear to be a definitive answer as of yet as to how many default components a body of knowledge of x has as well as how many total conceptual components can be in play

---

[33] To note, it is traditionally thought that working memory can contain seven elements or chunks plus or minus two, where elements are such various things as words and numbers. G. A. Miller, "The Magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review* 63: (1956), pp. 81-97. Recently, psychologist have discovered that the amount of chunks that working memory can simultaneously maintain depends on several variables such as the category of chunks used. For example, the capacity is seven chunks for digits and five for words. C. Hulme, S. Roodenrys, G. Brown, and R. Mercer, "The role of long-term memory mechanisms in memory span," *British Journal of Psychology* 86: (1995), pp. 527-536.

when we reason about x in working memory.[34] It may perhaps remotely be the case that working memory can handle enough components to account for super-structured hybrid concepts.

A second issue Weiskopf mentions is that there is no reason to believe that all the various bodies of knowledge are retrieved and used each time a person tokens a concept. Although different bodies of knowledge may work together in cognition, it appears that there are a relatively well-delimited number of types of bodies of knowledge of x that are used in specific acts of cognition. For example, in the previously and numerously discussed prototype studies in the third chapter such as for HIGHLY MORAL PERSON, there is no evidence that theories were present and qualifiedly used in working memory during these tests even though it was possible for participants to list a theory in the feature-listing task. Thus, it seems that strong pluralism is a more accurate description of the organization of our various bodies of knowledge as only a certain concept(s) of a category, rather than a whole super-structured concept, is brought forth into working memory for a particular higher cognitive competence act. It does not appear and there is no reason to believe that all of the bodies of knowledge of a particular moral category are in play for acts of moral cognition.

If the hybridist responds to these objections by claiming that only a part(s) of the hybrid concept may be recruited in working memory rather than the concept itself, then

---

[34] Machery, ibid., p. 11. Loyd Komatsu, "Recent views of conceptual structure," *Psychological bulletin* 112: (1992), pp. 520-521.

the hybrid view begins to collapse into strong pluralism. For, it is the strong pluralist rather than the hybridist who claims that only a single body of knowledge rather than all the bodies of knowledge can be brought into working memory. A single super-structured concept of a category is not always simultaneously recruited into working memory where concepts are assembled into complete thoughts as a hybridist traditionally claims.

While Weiskopf believes that his particular objections *begin* the collapse of hybridism into strong pluralism, we may perhaps complete the collapse by adding a further clarification to Weiskopf.[35] This clarification may also address the staunch hybridist who still insists that only parts of a super-structured concept need be recruited into working memory, and that therefore, there is still the possibility for hybridism. It also may address one who staunchly maintains weak pluralism, which contains a hybridism component. As concepts are theoretical psychological entities that are defined as the units of thought recruited from long term memory into working memory in order to play a specified causal role in higher cognition, when having thoughts about, for example, the morality of stopping global warming, one's MORAL concept is actively recruited into working memory to combine with other concepts to form the thought IT IS MORAL TO PREVENT GLOBAL WARMING. As discussed in the first chapter concerning the definition of concepts and at least dating back to Locke, concepts are understood as the individual units or constituents of thought that play a certain functional role in thinking and reasoning. Therefore, the MORAL concept is itself tokened in this instance.

---

[35] Special thanks to Weiskopf for help on this clarification.

This conclusion is the result from the clarification that concepts are the theoretical individual constituents of thought that are recruited into working memory to form complete thoughts. Given that this is the case, the hybridist is committed to the view that since the MORAL concept itself is brought into working memory in order to constitute the string of concepts that forms the complete thought and this MORAL concept is super-structured, all the relevant bodies of knowledge in the super-structured concept must be brought in as well. However, the concept MORAL is so loaded with various bodies of knowledge for the hybridist that it appears to lead to problems in that all the knowledge may not be able to be loaded into working memory given the generally understood small capacity size of working memory. Furthermore, it appears that all the various bodies of knowledge do not always simultaneously play a role in moral cognition. Thus, it seems to be the case that hybrid concepts in-and-of-themselves may be too large to play the relevant role in working memory, and the evidence suggests that when reasoning about ethics, prototype, exemplars, and theories do not simultaneously all play such a role in cases of moral cognition. If this is correct, this leads to the demise of the hybrid theory and the intermediate view, and rather, it supports strong moral concept pluralism.

6. Further Philosophical Implications

In this final chapter, I assess the further implications any of the previous chapters of the dissertation may have in ethics. The first implication is that the demise of the classical view means that many theorists in normative ethical theory and political philosophy must alter their approach in constructing their views. For, many of such theorists erroneously presuppose that our moral mental representations have classical structure. As another ramification, since our moral concepts do not have classical structure, we should expect there to be counterexamples to our moral concepts, but this does not necessarily mean that our concepts are wrong. Rather, we generally should adhere to those views that can best handle the most objections and the most serious counters.

A third implication is that the demise of the classical view and the existence of other accounts such as the prototype and exemplar theories provide a strong criticism against one of A.J. Ayer's arguments for emotivism. Finally, by relying on such things as studies in previous chapters that establish the viability of the prototype, exemplar, theory, and emotion theories, I argue against philosophers who use conceptual analysis in addressing the motivational judgment internalism/externalism and motivational Humeanism debates. I contend that such philosophers are not justified in reaching their conclusions on these debates.

6.1 The Demise of the Classical View and Theoretical & Applied Ethics

Stephen Stich, Alvin Goldman, and Mark Johnson have argued against the classical view of moral concepts by claiming the viability of the prototype theory along with its concomitant typicality effects.[1] While all three rely on typicality effects found in prototype studies to make their arguments against the classical view for moral concepts, Stich and Goldman provide no experimental evidence for the viability of the prototype theory for moral concepts. Thus, this conclusion against the classical view is in violation of PAC. Moreover, as we have seen in the third chapter, Johnson also fails to establish the viability of the prototype view for moral concepts. However, recall that I have shown in the third chapter that many moral concepts do have prototype structure which allowed me to argue against the classical view.

Although Stich, Goldman, and Johnson fail to discredit the classical theory, they do mention further ethical implications the demise of the classical theory may have on ethical theory. These three philosophers argue against normative ethical theories that require us to have mental representations with definitional structure. If an ethical theory claims that terms such as 'justice' have necessary and sufficient conditions and that we may arrive upon knowledge of such a definition, such a theory is fundamentally wrong in that our very concept JUSTICE does not have definitional structure. Given that ethical and political philosophies presuppose that we can fully mentally represent their respective normative categories, they need to take into account the findings in cognitive science in

---

[1] Johnson, ibid. Stich, ibid. Goldman, ibid.

order to construct a view that is at least psychologically plausible in the sense that such normative theories are required to match the pre-existing structure of our moral concepts. If one claims that normative theories can be mentally represented completely as such, then there must be a structural match. Now, some contemporary moral philosophers are immune from this criticism in that, for example, Richard Boyd does not posit necessary and sufficient conditions to his normative theory.[2] Boyd claims that there are a homeostatic cluster of properties for correct moral categories that are not necessary and sufficient conditions. While Boyd's view explicitly is one about correct moral categories, his theory implicitly maintains that we can mentally represent such categories. While for Boyd, mental concepts such as MORAL GOODNESS do refer to natural moral properties, as discussed in the first chapter, I will not dispute or contend for a particular theory of reference. Rather, here, we will merely assume conditionally for the moment that Boyd's theory of reference or content is correct since our primary reason for discussing Boyd in these contexts is not about whether his theory of content is correct. Rather, it is to show how if his theory of content is correct and moral concepts represent moral categories, his theory of moral categories, unlike many others, implies a corresponding psychologically plausible view of moral concepts.

---

[2] Recall from the fourth chapter that a homeostatic cluster of properties are a set of properties that are relevant to scientific induction and explanation and that are not necessary and sufficient conditions. Here, the presence of one or some properties is likely to lead to the presence of others within the set. Richard Boyd, "How To Be a Moral Realist," *Essays on Moral Realism,* ed. Geoffrey Sayre-McCord, (Ithaca: Cornell University Press, 1988), pp. 181-228.

In what he calls his 'Homeostatic Consequentialism,' he believes that moral goodness is a natural category that is defined by a cluster of mutually reinforcing properties such as the need for friendship, cooperation, and intellectual expression. What properties are human needs is a complex theoretical question that requires empirical evidence such as socio-political experimentation. For instance, we would not have understood the role political democracies play in the homeostasis of moral goodness without having the appropriate conditions that allowed for the first forms of limited democracy.

While Boyd believes the correct moral category of *moral goodness* is constituted by a cluster of properties concerning human need and the normative theory of consequentialism, when we mentally represent this moral category, it appears that we will do so using prototype and theory-theory structures. To note, although the folk may also have prototype and theory moral knowledge, such knowledge need not coincide with the perhaps more complex forms of prototype and theory knowledge had by the likes of Boyd. The cluster of properties of human need related to such things as friendship, cooperation, and intellectual expression will be mentally represented by prototypes, and consequentialism will be tokened by theory knowledge. As we empirically gain a better understanding of the cluster of properties, our moral mental representations develop and fill out as well. This kind of psychological concept acquisition is no different than, for example, empirically discovering a new natural kind such as a new species of animal and mentally representing this natural kind with new and developing prototypes as we learn

more about the animal. As we learn more about the new species, our concept of the new species likewise evolves to reflect the new knowledge gained. As we can see, this Homeostatic Consequentialism allows Boyd to escape the present criticism.

However, Plato (or Plato's Socrates) appears to be guilty of requiring us to have moral concepts with classical structure. Plato appears to hold that we can fully mentally represent the platonic form of the Good; a form that has necessary and sufficient conditions. Through examining and studying the nature of the Good, it is thought that we can somehow grasp implicit and largely tacit knowledge of what the Good entails in various moral circumstances. When a definition of the Good has been rejected by way of counterexample, we then continue on our quest for uncovering and specifying necessary and sufficient conditions by articulating new solutions that may avoid previous pitfalls.

For example, in the beginning of *The Republic*, Socrates argues with Cephalus and Polemarchus about what the definition of 'justice' or 'doing what is right' is. While Cephalus and Polemarchus lay out their definitions of 'doing what is right,' through the Socratic method, Socrates provides counterexamples to their definitions and thus, convinces them to alter their definitions as a result of the counters. This has the strong appearance that Socrates is presupposing that we can mentally represent 'justice' or 'doing what is right' with classical structure.

> 'That's fair enough, Cephalus,' I [(Socrates)] said. 'But are we really to say that doing right, consists simply and solely in truthfulness and returning anything we have borrowed? Are those not actions that can be sometimes right and sometimes wrong? For instance, if one borrowed a weapon from a friend who subsequently went out of his mind and then asked for it back, surely it would

> be generally agreed that one ought not to return it, and that it would not be right to do so, nor to consent to tell the strict truth to a madman?'
>
> 'That is true,' he [(Cephalus)] replied.
>
> 'Well then,' I said, 'telling the truth and returning what we have borrowed is not the definition of doing right.'[3]

Many normative ethical theories claim that moral categories have necessary and sufficient conditions and that we may mentally represent such categories. For instance, Kant appears to presuppose the viability of the classical view by claiming that the grounding of all our moral duties is based upon the categorical imperative, where the first formulation is a decision procedure to act only in accordance with that maxim through which you can at the same time will that it become a universal law.[4] Being allowed by a principle that one could rationally will to hold as a universal law is a necessary and sufficient condition for an act to be morally permissible. Jeremy Bentham appears to lay out necessary and sufficient conditions to utilitarianism by claiming that an action is right so long as it promotes the overall happiness or pleasure. Rosalind Hursthouse, a modern day virtue ethicist, claims that "[a]n action is right iff it is what a virtuous agent would characteristically (i.e. acting in character) do in the circumstances."[5] It also seems like Thomas Scanlon, a contractualist, argues that a necessary and sufficient condition for

---

[3] Plato, *The Republic, 2nd Edition*, trans. by Desmond Lee, (London: Penguin Books), 1974, p. 7.
[4] Immanuel Kant, *Groundwork of the Metaphysics of Morals*, Ed. by Mary Gregor, (Cambridge: Cambridge University Press), 1997.
[5] Her notion of "acting in character" is to act based on a firm, reliable, and stable disposition that leads towards moral excellence. Even if her theory relied on a definition that in some way deferred to prototypes or exemplars, she still in part posits a definitional structure which is not psychologically real. Rosalind Hursthouse, *On Virtue Ethics*, (Oxford: Oxford University Press), 1999, p. 28.

determining if an act is right is if it could be justified to others on grounds that they could not reasonably reject.[6]

Recently David Miller discusses numerous contemporary political philosophers' views, such as those by Ronald Dworkin and Thomas Nagel, concerning the boundaries within which the principles of distributive justice apply.[7] These views presuppose that there are necessary and sufficient conditions to such boundaries of justice and that we may mentally represent them, and Miller criticizes these views in that they are not successful necessary and sufficient conditions. For instance, Miller labels Dworkin and Nagel's view as the political coercion theory. This theory claims that the boundaries of justice are the boundaries of systems of legitimate political coercion. A system of legitimate political coercion is a system where individuals have certain freedoms restricted by laws and are subject to punishment if breaking such coercive laws. Moreover, there must be a legitimate justification for such laws, and the justification that succeeds is that the overall political system is distributively just. On this view, being the boundaries of a system of legitimate political coercion is a necessary and sufficient condition for being the boundaries of distributive justice, where an individual who falls outside such boundaries is an individual in which the full extent of the principles of distributive justice does not apply.

---

[6] To note, the necessary and sufficient condition of reasonable rejection for Scanlon is stated here in a simplified form. T. M. Scanlon, *What We Owe to Each Other*, Cambridge, MA: Harvard University Press, 1998.

[7] David Miller, "Justice and Boundaries," *Politics, Philosophy & Economics* 8: (2009), pp. 291-309.

Miller objects to the political coercion view on several accounts. As one objection, he argues that there are non-coercive settings, such as in clubs and churches, in which the principles of distributive justice do apply. For example, principles of equality for holding a certain office may come into play in some non-coercive groups. Hence, the boundaries of justice are not defined by the boundaries of systems of legitimate political coercion. Miller writes, "So although questions of distributive justice may become especially pressing between people whose relationships to one another involve routinely applied coercion, it seems that such relationships are neither strictly necessary nor sufficient to bring principles of distributive justice into play."[8] Regardless of whether Miller's objection is right that the principles of justice also apply in non-coercive settings, we may specifically criticize the likes of Miller, Dworkin, and Nagel for presupposing that that there may be classical structure to our moral concepts. The primary purpose of putting forth this example is to waylay this particular criticism against the above philosophers.

Recall from the third chapter that pervasive typicality effects were found for ethical concepts such as GOOD PERSON, JUST, BRAVE, and MORAL. Moreover, we examined prototype concept acquisition studies for RIGHT and WRONG ACTION. In this third chapter, from typicality effects found in prototype studies, *inter alia*, we argued that the classical view is not psychologically real for such moral concepts. Thus, while this ethical ramification or objection at hand fundamentally implicates many if not most

---

[8] Miller, 301.

normative ethical theories, philosophers must take alternative routes, such as Boyd, when constructing their views.

The Plato example points to a further ethical implication in that it is an open possibility that there will be counterexamples to our moral concepts given a lack of classical structure. This point is not stated by Stich, Goldman, or Johnson. Especially if we are assuming that correct moral categories must psychologically or structurally be able to be mentally represented completely as such by human minds, then simply denying someone else's moral conception based on one counterexample is too simplistic a move in attempting to refute this person's conception of, for example, JUST.[9]

At this point we will explore some cases of why a lack of classical structure will lead to the possibility of there being counterexamples. First, let us examine a concept that is thought to have classical structure: EVEN NUMBER = A NUMBER DIVISIBLE BY TWO WITHOUT REMAINDER. If I have this concept, in order for me to categorize a number as being even, it is necessary and sufficient for the number in question to be a number divisible by two without remainder. Without exception, only numbers that satisfy this condition can be classified as even numbers. There is a strict definition here with no room for counterexamples given the classical structure of my concept.

---

[9] This passage for the moment assumes that there are correct moral categories, but in this dissertation, I remain agnostic as to whether there are correct moral categories or not. Now, what would be an adequate refutation of someone's moral conception will be dependent on the exact structure of a correct category. For example, Boyd claims that there are a homeostatic cluster of properties for correct moral categories that are not necessary and sufficient conditions. One way to argue against his view would be to have a good number of counterexamples that precisely show that the cluster of properties can come apart in various ways.

Now, recall how we have seen a glimmer of why the demise of the classical view will allow for counterexamples when we examined Wittgenstein's prototype or family resemblance analysis of GAME in the third chapter. If a concept such as GAME has family resemblance rather than classical structure, then there can be counterexamples to parts of one's concept. For instance, if my GAME concept has prototype structure and is in part constituted by INCLUDES TWO OR MORE PEOPLE, then there can be a counterexample to this because I understand solitaire to be a game, but it only involves one person. The property of *including two or more people* may be a statistically frequent feature of games for me, but it is not a necessary condition.

For moral concepts, a lack of classical structure also will inevitably lead to the existence of counterexamples to analyses presenting necessary and sufficient conditions for those concepts. For example, my prototype of MORAL ACT may be in part constituted by DO NOT HIT OTHERS and DO NOT MAKE OTHERS FEEL BAD. Although these components play a role in influencing me to categorize certain acts, such as peaceful protests when people abstain from violence, as being moral acts, such constituents are not necessary conditions for me. For, I categorize acts where one physically defends oneself from an intruder in one's home as being morally permissible even though one strikes and harms the intruder. Moreover, they are not sufficient conditions because even if they are jointly satisfied, I still may not classify a particular act as being a moral one. For instance, if a couple did not even attempt to physically defend the lives of their children from a violent intruder in their home because they are radical pacifists, I would not

categorize this to be a moral act even though the couple did not attempt to strike the intruder and make the intruder feel bad. Rather, I perceive that there is some kind of moral failing in this circumstance.

Continuing with examples, my exemplar concept of MORAL ACT may be in part constituted by MY FRIEND NOT LYING TO THE POLICE OFFICER LAST YEAR and OBEYING THE LAW LAST NEW YEAR'S EVE. Recall from the second chapter that exemplars are themselves constituted by concepts that are about the generalizable moral attributes of the extension of the exemplars. For instance, MY FRIEND NOT LYING TO THE POLICE OFFICER LAST YEAR may be in part made up of TELL THE TRUTH. OBEYING THE LAW LAST NEW YEAR'S EVE may be in part constituted by BE A MODEL CITIZEN. TELL THE TRUTH and BE A MODEL CITIZEN may be generalizable and relevant to other different moral scenarios, such as when one takes the stand in a trial. However, TELL THE TRUTH and BE A MODEL CITIZEN are not necessary conditions for me because I would classify a person hiding Jewish persons up in his attic and lying to the Gestapo as being moral even though he lies to Nazi police officers and disobeys Nazi law by telling them that he is not hiding Jewish persons up in his attic. Furthermore, such components are not sufficient because even though particular situations may satisfy these represented conditions, I still may not categorize them as being moral acts. For instance, even though Albert did not lie to the officer and he was a model citizen by obeying the law and telling the truth to a Nazi officer that he is hiding Jewish persons up in his attic, I still do not classify his act as

being moral. Rather, I deem it to be an immoral act even though he did not lie to the officer and he obeyed the Nazi law.

As stated in the second chapter, the theory-theory for moral concepts claims that theories do not represent necessary and sufficient conditions, but rather, they occupy defeasible placeholder positions. Recall from the second chapter that by occupying such a position, theory knowledge can be understood as being generalizations that are taken to be defeasible based on, for example, the possibility of being defeated by new evidence or moral learning in the future. Hence, they are not consciously or subconsciously taken to be a strict definition. Insofar as this is the case, the theory-theory is very open to the possibility of there being counterexamples to generalized theories about what is morally right or wrong.[10]

In expecting counterexamples, we should adhere to those theories or views that currently can withstand the objections or are best protected against many to most counters, especially against those counters that are potentially the most damaging. Here, I allow for the possibility that one may hold a theory that is best protected from counterexamples but not from all of them because it may be possible that some individuals may have a defeated theory, but cannot think of an alternate theory at the

---

[10] While the classical theorist also may hold that the represented necessary and sufficient conditions can be open to revision due to error, they maintain that there is the possibility for the correct definition to be arrived upon, where such a definition is not defeasible. On the other hand, with the theory-theory, theory knowledge is always defeasible. Also, recall from the second chapter that if anything, theory knowledge may represent necessary but not sufficient conditions. This is another factor that differentiates the theory-theory from the classical view.

present moment that is undefeated. Rather, although their present theory is defeated from certain counterexamples, it is the best theory they can come up with at the present moment.

With all of the above in mind, as examples, if one is a rule-based deontologist, then one should construct a list of rules, that are mentally represented by prototypes, that can best handle the most objections and the most serious counters. Concerning the prototype view, successful counterexamples are not ones that show that the represented prototypical features simply fail to be effective necessary and sufficient conditions because such features are not even taken to be definitions in the first place. Successful counterexamples for the prototype theory are also not instances where the features of the counterexamples do not carry any weight at all because this is not to be expected. We should expect the qualities of a category to carry weight, where some are more important than others. Therefore, we should expect the features of the successful counterexamples to carry weight as well. This aspect of having weighted features is accurately mirrored in the prototype view. Rather, successful counterexamples for the prototype theory, for example, will be a group of important and central cases of good acts that do not match most of the prototypes of one's GOOD ACT concept and that do not match the heavily weighted prototypes either.

Moving on to the exemplar theory, if one is a virtue ethicist who believes that we should act as the virtuous agent would act, then one should focus on particular individuals, who are mentally represented by exemplars, who can best withstand a

thorough scrutiny of character. Recall that such exemplars themselves will be constituted by concepts that refer to things like more generalizable virtues such as honesty and benevolence. In deciding which individuals can best withstand a thorough scrutiny of character, one must look for individuals with certain generalizable attributes; attributes that can best handle the most relevant objections and the most serious counters. While keeping in mind that theory bodies of knowledge are defeasible, normative ethical views likewise should be adopted that presently are thought to withstand the current objections or most of the objections, especially the most damaging ones.

Finally, Goldman notes that Ayer in part argues for emotivism and for the view that moral judgments are the expression of our emotions by relying on the fact that we are unsuccessful at arriving upon a definition of our moral terms.[11] Ayer claims that since moral philosophers have great difficulty in providing a definition for moral terms, descriptivism in ethics should be replaced by emotivism. However, given that our moral concepts do not have classical structure, it should come as no surprise that it is hard to come up with a definition for moral terms. Moreover, given the viability of, for example, the prototype and exemplar theories for moral concepts, our concepts may still have descriptive content, although not a definitional descriptive content. In this manner, we can see that it does not immediately follow that descriptivism is false and therefore, emotivism is true just because our moral concepts do not have classical structure.

---

[11] Goldman, ibid. A.J. Ayer, *Language, Truth and Logic*, (London: Gollancz), 1936.

Since Stich, Goldman, and Johnson have not established their conclusions against definitionism for moral concepts, they cannot validate their further ethical ramifications. However, given that I have demonstrated in this dissertation that definitionism is not viable for many moral concepts, we finally may legitimately argue that their further ethical implications in ethical theory – concerning the objection against normative ethical views that imply that we can have psychologically unreal mental states and the contention against Ayer's particular argument for emotivism – hold.[12]

6.2 Types of Conceptual Analysis

We will proceed further in our inquiry of now examining and generally criticizing conceptual analysis arguments made for the motivational judgment internalism (MJI)/motivational judgment externalism (MJE) and motivational Humeanism debates in metaethical moral psychology. This will demonstrate the import of the previous chapters of the dissertation regarding the method of conceptual analysis for these two debates.

MJI will be understood to claim that there is a necessary connection between someone making a moral judgment and this speaker having some degree of motivation, whether that degree of motivation is defeasible or overriding. Some proponents of this view are Timmons and Gibbard.[13] Meanwhile, motivational judgment externalism

---

[12] John Jung Park, "Prototypes, Exemplars, and Theoretical & Applied Ethics," *Neuroethics*, forthcoming.
[13] Timmons, ibid. Gibbard, ibid.

(MJE), held by the likes of Brink and Shafer-Landau, denies this necessary connection.[14] Now, there are numerous other variations of internalist and externalist positions. For example, Michael Smith and Christine Korsgaard argue that when one makes a moral judgment, one is necessarily motivated unless one is practically irrational. [15] Insofar as such variations rely on the use of conceptual analysis, what I will later have to say about this debate will also apply to them. However, due to ease of linguistic expression, I will only discuss MJI and MJE as originally defined above, and I will not explicitly discuss alternate views such as Smith's and Korsgaard's, even though their theories are within the current purview.

For the final metaethical issue, motivational Humeans like Blackburn claim that beliefs by themselves do not motivate nor do they produce any motivational states.[16] Rather, it is necessarily the case that only emotions and other conative states motivate. On the other hand, anti-Humeans like Nagel argue that some beliefs by themselves are sufficient to motivate, or by themselves, some beliefs can lead to a motivational state.[17] Anti-Humeans deny the claim that only emotions and other conative states motivate.

Notice that both of these issues appear to importantly implicate concepts. As MJI is commonly and generally conceived, it claims that the connection between moral judgment and motivation must hold in virtue of the judgment itself, not in virtue of some

---

[14] Brink, Ibid. Shafer-Landau, ibid.
[15] Michael Smith, *The Moral Problem*, (Maleden, MA: Blackwell Publishing), 1994. Christine Korsgaard, "Skepticism About Practical Reason," *The Journal of Philosophy* (1986), p. 5-25.
[16] Blackburn, *Ruling Passions*.
[17] Thomas Nagel, *The Possibility of Altruism*, (Oxford: Clarendon Press), 1970.

contingent fact or reason.[18] Following Adina Roskies, let us call this *the intrinsicness requirement*.[19] We may therefore interpret the intrinsicness requirement for MJI as claiming that the moral concept in a judgment is at least in part constituted by a motivational state. If the concept is so constituted, then this satisfies the intrinsicness requirement because the motivational state is what at least in part constitutes the moral concept, where the moral concept in turn is what constitutes the judgment. In such a circumstance, motivation stems from the judgment itself. Therefore, in order to see whether MJI is right or not, we should examine whether it is necessary that all moral concepts contained within sincere moral judgments are at least in part constituted by motivational states. Finally, motivational Humeanism entails that the relevant concepts in motivating judgments must be at least in part constitutively linked to some conative state, while anti-Humeanism entails that at times the relevant concepts in motivating judgments or judgments that lead to motivation are not constitutively linked to some conative state.

Before criticizing the use of conceptual analysis in the MJI/MJE and motivational Humeanism debates, first we need to explore the various kinds of conceptual analysis as this is a very common method used by philosophers in addressing these two issues. However, before classifying the various kinds of conceptual analysis, we will initially have to discuss several important distinctions. First we must invoke the

---

[18] Roskies, ibid. Thomas Nagel, ibid.
[19] Roskies, ibid.

concept/conception distinction from the first chapter. We may assume that this distinction is in play for the remainder of the dissertation. As we shall see, there are various kinds of conceptual analysis, and we will need this distinction in order to articulate the differences between some of them. Recall that a concept is an abstract entity that is metaphysical in the sense that it is a correct concept whose components represent the correct classificatory descriptions. Meanwhile, a conception is a mental representation in the head of an agent that is epistemic in the sense that its components represent the properties we actually think about when we categorize or reason about the extension of the conception regardless of whether such properties are correct or incorrect classifiers. Concepts and conceptions will both appear in all capitalized letters.

Although we have partially discussed it earlier in this chapter with Plato, generally speaking, conceptual analysis is picking a concept or conception and attempting to uncover or unpack the constituent components of the concept or conception through use of non-empirical and non-experimental a priori intuitions that are drawn from thought experiments and at times real life cases.[20] Notice that I have first detailed conceptual analysis as being either about uncovering abstract concepts or mental representations. As we shall see, some understand it to be about concepts, but others view it to be about conceptions.

---

[20] There is an empirically-oriented way of analyzing a conception. An analysis of the explanatory or theoretical role that a conception plays in science might support one analysis rather than another. However, for our purposes, the metaethicists at hand engage in conceptual analysis in an a priori way.

Also, philosophers can understand 'intuition' in conceptual analysis to mean different things. Different philosophers may attach different psychological or epistemological qualities to what an intuition is. For instance, Hugo Mercier and Dan Sperber describe an intuition as produced by "processes that take place inside individuals without being controlled by them."[21] On this account, the spontaneous inferences that justify an intuition are not consciously accessible to the agent. Jeff McMahan simply describes a conceptual analysis intuition in ethics as a strong spontaneous moral judgment.[22] Colin McGinn claims that they are "one's considered judgments about actual and possible cases."[23] Still, others may understand intuition in the standard epistemic sense, such as with G. E. Moore and more recently Robert Audi, where it is a non-inferentially justified belief that may be arrived upon through careful and perhaps long reflection.[24] While there are various ways of understanding what an intuition is when, for example, from a thought experiment a philosopher says that it is simply intuitively obvious that the intuition is correct, my use of the term will be a use that could be taken in any of the above senses. As we proceed, I do not believe that this will be detrimental to our final aims.

---

[21] Hugo Mercier and Dan Sperber, "Intuitive and reflective inferences," *In two minds: Dual processes and beyond*, ed. by J. Evans & K. Frankish, (Oxford: Oxford University Press), 2009.
[22] Jeff McMahan, "Moral Intuition," *Blackwell Guide to Ethical Theory*, ed. by Hugh LaFollette, (Oxford: Blackwell), 2000.
[23] Colin McGinn, *Truth By Analysis*, (Oxford: Oxford University Press), 2012.
[24] G. E. Moore, *Principia Ethica*, 5th edition, (Cambridge: Cambridge University Press), 1968. Robert Audi, *The Good in the Right*, (Princeton: Princeton University Press), 2004.

Another distinction will need to be made between what I call the *formal* versus *material* aspect of conceptual analysis. This will aid us in our future discussion of conceptual analysis. Here, I will elaborate upon this distinction in terms of the conceptual analysis of conceptions, but as we shall later see, this distinction can also be thought to hold for the domain of concepts, *mutatis mutandis*. Moreover, I take this distinction to be pertinent to the use of conceptual analysis specifically in the MJI/MJE and motivational Humeanism debates, and this is all we need for our purposes. I make no strong claim here as to whether this distinction holds for conceptual analysis in all other various types of fields of inquiry. At this point, we will now move into the discussion of the conceptual analysis of conceptions both in light of the formal/material distinction and in terms of a general discussion of what the conceptual analysis of conceptions is.

The formal aspect is the analysis of the components of the conception in question irrespective of whether the propositional content of the components could be realized. The MJI/MJE debate is often argued on the a priori grounds of whether the existence of the amoralist is conceivable. Some philosophers have attempted to argue against MJI by introducing the notion of the amoralist. For example, David Brink argues that the MJI advocate fails to take the amoralist seriously, where the amoralist may make sincere moral demands but still remains unmoved.[25] In this case, the formal aspect of conceptual analysis pertains to the actual analysis of the conception AMORALIST or perhaps MORAL

---

[25] David Brink, *Moral Realism and the Foundation of Ethics*, (Cambridge: Cambridge University Press), 1989.

JUDGMENT. By analyzing our intuitions on these conceptions, we can perhaps figure out, for example, whether our AMORALIST conception is constituted by POSSIBLY COULD EXIST. We may also discern whether, for instance, our MORAL JUDGMENT conception is constituted by NECESSARILY HAVING SOME DEGREE OF MOTIVATION. However, the material aspect is concerned with whether the propositional content of the intuition belief in the formal aspect can be said to be true. While the formal element is, for instance, about the a priori analysis of the conception AMORALIST and what constitutes this mental conception, where an intuition belief may be AMORALISTS POSSIBLY COULD EXIST, the material element takes the propositional content of this intuition belief – namely, 'Amoralists possibly could exist' – and is interested in whether the proposition is true. Restated in perhaps more clear terms, the formal element is the analysis of the conception's components on the basis of a priori armchair investigation, and the material element is concerned with the separate question of whether the conception is true in all possible worlds. Another way to look at it is that formally, one's conception MORAL JUDGMENT may be constituted by NECESSARILY HAVING SOME DEGREE OF MOTIVATION. Here, motivation is merely being *purported* by the intuition to at least in part constitute all speakers' actual moral judgments. However, the material element is asking whether it is the case in all possible worlds that when all people make sincere moral judgments, all such judgments are actually at least in part constituted by motivational mental states, where it is truly the case that such judgments actually involve the speaker having the relevant motivation. For example, if Lucy is a real person in some possible world such as

the actual world and she forms the judgment STEALING IS WRONG, is it the case that in this world her conception WRONG in this judgment is actually at least in part constituted by, for instance, a motivational state? This pertains to the material element. Notice that while the formal aspect in our example is concerned with the conception MORAL JUDGMENT and its constituents in the head, the material aspect in our same example is rather about the conception WRONG and its constituents.

In regards to the conceptual analysis of conceptions, David Brink understands MJI to be a modal claim:

> Internalism (of any form) has at least three distinguishable components. The first claim is that moral considerations *necessarily* motivate or provide reason for action. The second and third claims come out of the internalist thesis that it is the concept of morality that establishes this. Since it is the *concept* of morality that shows that moral considerations necessarily motivate or provide reasons for action, this claim about the motivational power or rationality of morality must be a priori. Since it is the *concept of morality* that determines this fact, the rationality or motivational power of moral considerations cannot depend on substantive considerations such as what the content of morality turns out to be, facts about agents, or the content of the correct theory of rationality.[26]

The belief that MJI is a modal claim should not be surprising since MJI and MJE explicitly propound material modal conclusions. MJI advocates say that there is a necessary connection between making a judgment and being motivated to some degree, while MJE advocates deny the necessary connection. Moreover, MJI and MJE theorists take their respective material modal conclusions to be true.

---

[26] Brink, Ibid., p. 42.

For the motivational Humeanism debate, Thomas Nagel argues through a priori thought experiments that individuals who have prudential practical reasons for performing some action in the future may be motivated in the present by such reasons rather than from a present desire to satisfy future interests. However, he takes the formal element of conceptual analysis to not only be a conclusion regarding the constituents of his conception and what he conceives, but he also believes it entails material conclusions about all possible worlds which includes the actual world and actual human nature. Nagel takes the propositional content of his intuition to be true. For instance, in discussing his anti-Humeanism, he states, "[M]y suggestion sheds more light on the operation of prudence and on human nature in general."[27] This is consistent with the fact that motivational Humeans and anti-Humeans tend to state their claims in material modal terms. Humeans tend to claim that necessarily only conative states motivate, while anti-Humeans tend to deny this necessary claim.[28] Humeans and anti-Humeans understand their respective claims to be true.

---

[27] Thomas Nagel, *The Possibility of Altruism*, (Princeton: Princeton University Press), 1970, p. 38.
[28] I say 'tend to' because admittedly, it is not exactly clear in the literature whether the motivational Humeanism debate is a conceptual modal claim or whether it is an empirical issue about the actual world. Notice in the above Nagel example on his anti-Humeanism, he appears to make a conceptual modal claim. However, Shafer-Landau in certain (but not all) cases appears to understand it to be an empirical claim, where he attempts to justify anti-Humeanism by relying on his own psychological phenomenological experiences of acting. However, he also claims that Humeans believe that "desires are necessary for motivation (p. 122)." While I discuss it as a conceptual modal claim, if one understands it to be an empirical claim about the actual world and not a modal claim, then the lack of empirical confirmation and PAC directly can be used to criticize most of those who make such a strong claim on this debate since most do not use accurate experimental evidence on the debate in question. Shafer-Landau, ibid.

In regards to the conceptual analysis of conceptions for the MJI/MJE and motivational Humeanism debates, it is generally thought that there is a link between conceivability, metaphysical possibility, and a priori necessity. For instance, through conceptual analysis of conceptions, if we cannot conceptually conceive of someone making a moral judgment without being motivated in some way because conceiving otherwise would lead to a contradiction, then it is an a priori necessity that there is a connection between moral judgment and some degree of motivation. This a priori necessity apparently holds in all possible worlds. As we can see, conceptual analysts of conceptions for the MJI/MJE and motivational Humeanism debates believe that there is a link between the formal and material aspect, where a priori intuitions not only show us what constitutes our formal conceptions and tells us what is and is not conceivable, but such intuitions also can entail relevant material conclusions about the truth or falsity of the propositional content of the intuition beliefs via the conceivability-metaphysical possibility-a priori necessity link. [29]

Moving along with further distinctions, I take it that the conceptual analysis of conceptions is taken by many to be about the analysis of folk conceptions. For instance, in defending this brand of conceptual analysis, Frank Jackson states:

---

[29] If a philosopher has conception C of phenomenon P, then the philosopher will believe C of P. However, that does not show that the philosopher has the meta-view that, because C is her conception of P, C is necessarily true (qua truth with a capital T) of P. For, the philosopher may hold a deflationary theory of truth. Concerning the MJI/MJE and motivational Humeanism debates, most metaethicists involved in conceptual analysis believe that their relevant conclusions on such issues are necessarily true and not merely true in the deflationary sense.

> But how should we identify our ordinary conception? The only possible answer, I think, is by appeal to what seems to us most obvious and central about free action, determinism, belief, or whatever, as revealed by our intuitions about possible cases. Intuitions about how various cases, including various merely possible cases, are correctly described in terms of free action, determinism, and belief are precisely what reveal our ordinary conceptions of free action, determinism, and belief, or, as it is often put nowadays, our folk theory of them.[30]

Likewise, concerning his use of conceptual analysis on a variety of moral issues such as on motivational Humeanism, Smith writes:

> To say that we can analyse moral concepts, like the concept of being right, is to say that we can specify which property the property of being right is by reference to platitudes about rightness: that is, by reference to descriptions of the inferential and judgmental dispositions of those who have mastery of the term 'rightness.'[31]

Usually a telltale sign that a philosopher understands conceptual analysis to be about coming to a clear understanding of folk conceptions is when authors offer their intuitions by opening with, "*We* wouldn't say that…" or something to that effect. For instance, Shafer-Landau's chapter on MJI is littered with phrases such as "*We* can apparently imagine…"[32] and "*we* can imagine an agent…"[33] In relation to the same metaethical issue, Brink's writing is suffused with such passages. In one selected paragraph, we have phrasings such as, "*We* can imagine someone…", "*We* may even think that such a person is merely possible…", "But *we* do think that such a person is

---

[30] Frank Jackson, *From Metaphysics to Ethics: A Defense of Conceptual Analysis*, (Oxford: Oxford University Press), 1998, p. 31.
[31] Smith, ibid., p. 39.
[32] My italics. Shafer-Landau, ibid., p. 151.
[33] My italics. Shafer-Landau, ibid., p. 150.

merely possible…", etc.[34] Due to such reasons, although they do not explicitly state their view on conceptual analysis and it may not be wholly clear on what they view conceptual analysis to be, I understand the likes of Shafer-Landau and Brink to maintain folk understandings of the conceptual analysis of conceptions.[35]

However, it must be noted that one may take the conceptual analysis of conceptions to be the analysis of *expert's* conceptions. For instance, McGinn claims that all of philosophy in the present as well as in history is fundamentally the a priori search for essences by examining experts' psychological conceptions. When experts examine their own conceptions a priori, they gain access to the essential truths and nature of the extra-mental real world.[36] For, philosophy's inherent goal is to arrive upon conclusions of the nature of reality, not just merely the structure of our formal conceptions. This is best done by introspectively trained philosophers who examine only their own conceptions.[37] Given that this may be the case, in the conception domain, we have the possible demarcations of *folk conceptual analysis of conceptions*, which analyzes folk

---

[34] My italics. Brink, ibid., 48.

[35] Even if they really espouse some other form of conceptual analysis, I eventually will provide criticisms against all forms of conceptual analysis.

[36] To note, McGinn explains the connection between mind and world by arguing for a teleosemantics view. In this fashion, our conceptions refer to things in the world so long as our conceptions have the function of representing such things, and this provides the link between mind and world.

[37] For further philosophers who hold that it is expert's intuitions that only matter, see T. Williamson, "Armchair philosophy, metaphysical modality and counterfactual thinking," *Proceedings of the Aristotelian Society*, 105: (2005), p. 1-23. S. D. Hales, *Relativism and the Foundations of Philosophy*, (Cambridge, MA: MIT Press), 2006. K. Ludwig, "The epistemology of thought experiments: First person versus third person approaches," *Midwest Studies in Philosophy*, 31: (2007), pp. 128-159.

conceptions, and *expert conceptual analysis of conceptions*, which analyzes philosopher's conceptions only.

Now, we will turn our attention to describing the conceptual analysis of concepts. First, we will discuss it in light of the formal and material distinction. For, this distinction can be thought to also hold for the conceptual analysis of concepts. For example, if I attempt to analyze a priori the correct metaphysical concept MORAL JUDGMENT, my intuitions may come to the conclusion that this abstract concept is constituted by NECESSARILY HAVING SOME DEGREE OF MOTIVATION, where motivation is being purported to at least in part constitute all speakers' moral judgments. The formal element is what the philosopher a priori understands the constituent of the concept to be. The material element is then concerned with, for example, the truth of the following claim that 'in all possible worlds, all moral conceptions in sincere moral judgments in all moral agents is at least in part constituted by a motivational mental state,' where it actually is the case that some motivational state really at least in part constitutes everyone's sincere moral judgments. The formal element is the analysis of the concept's components on the basis of a priori armchair investigation, and the material element is concerned with the separate question of whether the concept constituents are true in all possible worlds.

Philosophers believe there is a strong link between the formal and material elements of the conceptual analysis of concepts. For, they are analyzing the *correct metaphysical* concept, and any constituents of the concept are taken to be correct or true.

Hence, if a philosopher engaged in the conceptual analysis of concepts a priori takes the concept MORAL JUDGMENT to be constituted by NECESSARILY HAVING SOME DEGREE OF MOTIVATION, then this philosopher believes that it is true that there is a necessary connection between making a moral judgment and having some degree of motivation in virtue of what the philosopher thinks the correct concept is. This is consistent with the general claims made by philosophers in the MJI/MJE and motivational Humeanism debates because they draw material modal conclusions; conclusions which they understand to be true. Philosophers in these debates that espouse the methodology of the conceptual analysis of concepts can still draw their modal conclusions because there is an apparent link between the formal and material elements.

For example, Georges Rey, in a well-known article concerning the concept/conception distinction, takes conceptual analysis to be the conceptual analysis of concepts, where an a priori conclusion entails that the propositional content of the intuition is true.[38] For Rey, the formal element is the analysis of the abstract concept, while the material element is concerned with whether the intuition arrived upon in the formal element is true. In analyzing abstract concepts and their constituents, we then can make relevant claims about the truth of the relevant proposition.[39] For instance, Rey

---

[38] Rey, ibid.

[39] The ontological difference between concepts and conceptions must be noted because the constituents of abstract concepts is not directly epistemic in the way that the constituents of conceptions are epistemically already in the head. Conception constituents are what we can a priori conceive. In this respect, the conceptual analysis of conceptions is about how we conceive of possible worlds, and therefore, there is a link between such a priori analysis, possibility, a

writes, "Concepts, that is, would seem – at least in the first instance – to be about the world and how *it* divides up; not about how we might divide up our methods of investigating it."[40] He also states in distinguishing concepts from conceptions:

> The division I draw between these remaining functions corresponds to a crucial, if very battered distinction in philosophy between *metaphysics* and *epistemology*, or between issues surrounding *how the world is* (what exists, what is true) and issues surrounding *how we know, believe, infer, how the world is*.[41]

I will understand the conceptual analysis of concepts to generally claim that such analysis should be done by those professional philosophers who have read the literature on the subject matter, are well trained in logical thinking and getting in touch with abstract entities, etc. In a word, since what is at hand is the analysis of abstract entities or concepts, similar to advanced mathematics and logic where it mostly is the professional mathematician or logician's job to conduct a priori analysis on concepts, the conceptual analysis of concepts will be understood to primarily fall within the domain of those trained metaphysicians who understand conceptual analysis to be the analysis of abstract

---

priori necessity, and thus, the material element. However, since the conceptual analysis of concepts is not directly epistemic, it does not appear to deal with conceivability in the same way as with the conceptual analysis of conceptions. Thus, for the conceptual analysis of concepts, philosophers' a priori intuitions supposedly are directly channeled to what the constituents of the relevant metaphysically correct concept are. It appears for the conceptual analysis of concepts that there is no conceivability-possibility-necessity link. Rather, intuitions about concepts supposedly are directly linked to what is purportedly true.

[40] Rey, ibid., p. 289.

[41] Rey, ibid., p. 284.

concepts. Here, Rey claims that experts may have access to concepts through their own a priori intuitions.

6.3 On the Use of Conceptual Analysis for MJI/MJE & Motivational Humeanism

It is time now to attempt to criticize the various forms of conceptual analysis in light of the MJI/MJE and motivational Humeanism debates. First, we will discuss expert and folk conceptual analysis of conceptions. A criticism we may lay against expert and folk conceptual analysis of conceptions is that empirical work is relevant and required to assess the claims about the nature of the constituents of the conceptions that are being analyzed. Empirical work is relevant to the formal element. For, conceptions are mental representations in the head, and I have shown throughout this dissertation that, although theoretical work is required in order to frame empirical investigation, the examination of the nature of one's moral conceptions is importantly an empirical affair. Thus, for instance, whether one wants to know what constitutes the expert's AMORALIST conception or the folk's AMORLIST conception, one needs empirical evidence to secure a conclusion on this matter.

One may object to this by stating that philosophers and/or the folk do have clear access to their conceptions, so no experimental evidence is required to figure out what constitutes their conceptions in the formal aspect. As the objection goes, we may view the entire existence of the relevant areas of metaethics as being a kind of exploration of our conceptions. While there is no evidence from non-theorists and while the conditions

have not been precisely controlled, we still can be said to have a lot of evidence about the kinds of conceptions people possess. Nevertheless, especially given reasons that not everything about our inner minds and deepest thoughts may be consciously accessible to us through introspection and that PAC requires us to have experimental evidence on the abstract conception in question to draw constitution claims on the structure and knowledge that make up the conception, we need to run empirical studies on the relevant conceptions.[42] Moreover, another reason why we need to run experiments is that I have shown in previous chapters that many of our moral conceptions can be constituted by prototype, exemplar, theory, or emotion bodies of knowledge. This allows for possible variation in philosophers' and the folk's moral conceptions since there is the possibility that individuals have different bodies of knowledge for a conception and even if two individuals have identical kinds of bodies of knowledge, such as prototypes, there still may be great variance in their respective prototypes.[43] Recall the Smith, et al. study in

---

[42] Those philosophers who espouse an experts-based view may claim that philosophers have a more clear access to the nature of their moral conceptions than the folk, so they are immune to this objection. However, we will examine a study below that shows that professional philosophers are more susceptible than folk to ordering effects and the concomitant grasping of their proper intuition beliefs. This places a burden of proof on those relevant philosophers to empirically justify their claim that philosophers have a more clear and undisturbed access to the relevant conceptual constituents.

[43] One may wonder whether in such a case two individuals have different conceptions. This issue of conception individuation will be discussed shortly. It later will be shown how if philosophers' or folk's conceptions are interpersonally different (within their own respective domains), then there can be no convergence in philosophers' or folk's intuitions, respectively. This is highly problematic for conceptual analysis of conceptions because this general view predicts that there will be convergence in philosophers' or folk's intuitions. Therefore, for the sake of argument, I presuppose here that philosophers' or folk's conceptions are not different. However, even if this

the third chapter in which they found that the constituents in people's GOOD PERSON prototype conception substantially varied across seven different cultures. Also, based on the metaethical literature, there appears to be substantial variance amongst philosophers' intuitions on the MJI/MJE and motivational Humeanism debates. The relevant conceptions such as AMORALIST and MORAL JUDGMENT have the appearances of being something like Gallie's notion of an 'essentially contested conception.'[44] Gallie claimed that there are conceptions whose use leads to endless disputes. Users of the conception will give different weights and interpretations to the constituents of the conception, and they will provide arguments for their particular understanding while fully knowing that others hold incompatible views concerning the extension of the conception. The conceptions relevant to our two metaethical debates at least appear to be candidates for being essentially contested conceptions in that the variance in philosophers' intuitions has existed for hundreds upon hundreds of years. In fact, David Bourget and David Chalmers have run a survey asking professional philosophers what their views are on a host of philosophical issues. They found that 35% of the sample population believes in MJI,

---

is the case, it will be shown that there still appears to be no convergence in intuitions, and therefore, as presently stated, experiments need to be run on conceptions. In this respects, one may think of this point about variance I make as being conditional on if the relevant conceptions are not different.

[44] Walter Bryce Gallie, "Essentially Contested Concepts," *Proceedings of the Aristotelian Society*, Vol.56: (1956), pp.167-198.

30% hold MJE, and 35% maintain some other view.[45] All of the above variance suggests that there may be substantial divergence in the constituents of our moral conceptions, where conceptual analysis assumes that the opposite is the case, where there will be uniformity in philosophers' or the folk's conception constituents over time. Due to this conflict and the worry that it produces, this provides further reason to maintain that experimental studies need to be run on one's conceptions and a specified group of individuals' conceptions in order to sufficiently justify a moral conceptual analysis of conceptions claim. Intricate, subtle, and nuanced experimental tests about the constitution of our conceptions helps us to gain better access to what our conceptions contain.

Now, it is just such evidence concerning the formal element of the conceptions being analyzed that has largely and generally been absent in the MJI/MJE and motivational Humeanism literature for those who espouse any of the conceptual analysis of conceptions views. For instance, MJI advocates such as Timmons, Mackie, Hare, and Gibbard do not utilize such data.[46] Externalists such as Brink and Shafer-Landau do not use the proper evidence when propounding their claims either.[47] Motivational Humeans such as Blackburn and Smith and anti-Humeans like Nagel, McDowell, Shafer-Landau,

---

[45] David Bourget and David Chalmers, "What do Philosopher's Believe?" *Philosophical Studies*, forthcoming.

[46] Mackie's internalism can be viewed as a kind of error-theory internalism. Timmons, ibid. John Mackie, ibid. R. M. Hare, ibid. Allan Gibbard, ibid.

[47] Brink, ibid. Shafer-Landau, ibid.

Dancy, and Scanlon also appear to not provide the requisite support.[48] Due to this fact, their claims are unwarranted and tests need to be run either on expert or folk conceptions. Therefore, at best, an agnosticism is warranted on the MJI/MJE and motivational Humeanism debates in light of the conceptual analysis of conceptions.

Shifting now to the conceptual analysis of concepts, as an example case, Plato's Socrates may be interpreted as an instance of one who has engaged (or rather created) the MJI/MJE debate by relying on the conceptual analysis of concepts methodology.[49] In the *Protagoras*, Socrates is attempting to come upon a definition of the Good with Prodicus, Protagoras, and Hippias. It appears that Socrates claims that it is part of the form or the concept of the Good that MJI is true. For instance, Plato's Socrates says, "Then if the pleasant is the good, no one who knows or believes there is something else better than what he is doing, something possible, will go on doing what he had been doing when he could be doing what is better."[50]

I will now examine potential objections against the conceptual analysis of concepts that also applies to expert conceptual analysis of conceptions in virtue of the fact that both of these views claim that we must rely on the intuitions of the expert

---

[48] It is not completely clear to me as to whether McDowell, Dancy, and Scanlon adhere to a conception form of conceptual analysis. Even though I list them as being such, even if they adhere to a concept version of conceptual analysis, I still lay criticisms against such methodologies later as well. Simon Blackburn, ibid. Smith, ibid. Nagel, ibid. McDowell, ibid. Shafer-Landau, ibid. Jonathan Dancy, ibid. Scanlon, ibid.

[49] For our purposes, it matters not whether Plato's Socrates is expressing Plato's own views or Socrates'.

[50] Plato, *Protagoras*, ibid.

philosopher. To note, one way to tell the difference between these two types of conceptual analysts is to know if a particular analyst is a Platonist that believes the relevant categories are abstracta that we can get in touch with using our a priori intuitions. Such a person is an expert conceptual analyst of concepts. On the other hand, an expert conceptual analyst who does not believe that the relevant categories are abstract objects most likely maintains an expert conceptual analysis of conceptions view. First, for those who believe that philosophers are better at drawing correct intuitions on abstract concepts or on their own experts' conceptions as compared to the folk's intuitions on abstract concepts or on folk's own conceptions, respectively, no evidence has been given to demonstrate that this is so. Rather, philosophers such as McGinn merely just assume this to be correct. However, it appears that the burden of proof is on such adherents to an experts-based view of conceptual analysis to actually provide concrete evidence for their claim to superiority in a non-question begging way since their positive philosophical methodology is so dependent on the notion that experts' intuitions are better.[51] In other words, it appears that it is an empirical matter whether philosophers are better at the a priori analysis of concepts or their own experts' conceptions than non-expert competent speakers of a language.[52] Once the correct criterion has been laid down for what the

---

[51] Kristopher Ahlstrom-Vij, "Truth by Analysis: Games, Names, and Philosophy," *Notre Dame Philosophical Reviews*, 2012.06.05.

[52] While this may presuppose that philosophers and the folk do not have different conceptions, we will discuss the subject matter of conception individuation shortly. What we will later find is that if philosophers' and folk's conceptions are different, then this is problematic for the expert conceptual analyst because this will not allow for philosophers' conceptual analysis to be better

correct concept or conception is, tests must be run in order to sufficiently see whether philosophers or experts are better at the analysis of the nature of the supposed correct concept or conception. Of course, this assumes that the relevant inquiry will have been completed and a definitive conclusion will have been reached. Assuming that this is the case, it behooves philosophers such as McGinn to then run studies to see whether or not the folk are overall better or worse at the relevant conceptual analysis of conceptions or concepts. For, even if there already is a general consensus on the correct conclusion from experts, it still may turn out that the folk are better and have a stronger consensus on the correct conclusion. The burden falls on the philosopher who makes the positive contention that expert's intuitions are superior to actually prove that expert's intuitions are superior. Their conclusion must be justified, and such justification requires empirical evidence. For example, if one has in hand the supposed correct prototype for the conception AMORALIST, then one may run a prototype feature-listing task, as discussed in detail in the second chapter, on the folk and professional philosophers to see what their prototype conception of AMORALIST is. One may then run a similarity score measure on how similar the folk and experts' conceptions are to the purported correct prototype. Once this has been done, whichever group has the higher similarity score is the group that

---

than the folk's conceptual analysis. Therefore, for the sake of argument, I here presuppose that philosophers' and folk's conceptions are not different. However, it will be shown how by not being different, philosophers and the folk can still potentially have contradictory intuitions, and as stated, empirical studies still need to be run. In this respects, the present objection may be thought of as being conditional on if the relevant conceptions are not different.

is better at conceptual analysis. Given that no such tests have been run to prove that expert's intuitions are better than the folk, one is not justified in concluding that expert's intuitions are superior. As the aims of this chapter are to explore further philosophical implications my pluralistic theory of concepts has in ethics, the objections provided thus far give some good reason to believe that metaethical claims in the MJI/MJE and motivational Humeanism debates that are based on the a priori conceptual analysis of concepts and expert and folk conceptual analysis of conceptions are not justified. We now proceed by further examining possible objections against conceptual analysis for our two metaethical debates.

A potential objection against expert conceptual analysis of conceptions is that given that people may have non-identical kinds of knowledge constitute their conceptions and even though they may have identical kinds of knowledge, the conceptual constituents still may significantly diverge, it may be the case that experts and the folk have different moral conceptions from each other, where the expert and folk conceptions are not the same. In claiming that experts' conceptual analysis of their own conceptions is superior to folk's conceptual analysis of their own folk conceptions, experts assume that experts' and folk's conceptions are the same or not different such that experts can make such a claim to superiority. If they have individuated or different conceptions, then one cannot claim that philosophers are better than the folk at conceptual analysis of conceptions because like apples and oranges, their different conceptions and therefore, their respective analysis of them cannot be properly compared to one another such that one

may say that one's conception and analysis of one's conception is better or more accurate than the other's.[53] If two people have different conceptions, then this means such an above specified comparison is not possible, where intuition beliefs that are made up of different conceptions have different truth conditions. Although later I will explain why I do not adopt this potential objection so nothing substantial hangs on our discussion of it, this potential objection at hand against expert conceptual analysis of conceptions presupposes that it is possible that the philosophers' and folk's conceptions may be different such that the above comparison cannot be made; a presumption and potential objection that we will now further entertain and explore. To note, within the present context, to have "different" or individuated conceptions is a heavily theory-laden notion, and we will now further discuss and further elaborate upon the meaning of these exact terms of having 'different or individuated conceptions' that will apply to the rest of this section. As we shall see, having different conceptions does not necessarily mean that we simply have non-identical constituents to our conceptions.

Now, philosophers may decide to differentiate conceptions in various ways. How philosophers decide to differentiate conceptions is relevant to the present entertained objection because one's theory of conception individuation may determine whether or not the present potential objection is correct that the philosophers and the folk have different conceptions. There are two major general theories of conception individuation, where

---

[53] Notice that the door is still open here for some level of interpersonal communication such as between two psychologists discussing what constitutes their respective different conceptions by discussing psychology tests that have been run on their relevant conceptions.

each theory presents a different meaning to 'conception individuation' and what it is to have different conceptions such that the different conceptions cannot be so compared to one another as discussed above. To note, here I merely describe both theories of when conceptions are different, but I do not critically assess them due to reasons to be elaborated upon shortly concerning the fact that I take no stand in this dissertation on a theory of conception individuation. For the first view of conception individuation, some like Ned Block, Prinz, and Paul Churchland claim that conceptions can be differentiated from each other by cognitive content (sense) and reference.[54] Let us call this specific theory of individuation *the internalist view*.[55] They claim that if there is some degree of internal psychological similarity in the constituents of both our MORAL conceptions and our MORAL conceptions co-refer, then our conceptions can be said to be the same conceptions and are not different from one another.[56] On this theory, what it means to have two conceptions that are not different from each other is to have two conceptions that share the same reference and have some degree of similarity in cognitive content. Remember again, even though one may have objections to this internalist view, I do not

---

[54] Ned Block, "Advertisement for a semantics for psychology," *Midwest Studies in Philosophy Vol 10: Studies in the Philosophy of Mind*, ed. by P. French, T. Uehlin, H. Wettstein, Minneapolis: University of Minnesota Press, 1986, pp. 615-678. Ned Block, "Functional Role and Truth Conditions," *Proceedings of the Aristotelian Society,* LXI, 1987, pp. 157-181. Prinz, *Furnishing the Mind.* Paul Churchland, *Plato's Camera*, (Cambridge, MA: The MIT Press), 2012.

[55] To note, the internalism/externalism distinction at hand for conception individuation is not necessarily meant to mirror the internalism/externalism distinction for theories of content.

[56] What degree of posited similarity that is required in order to have the same conception can vary depending on the theorist. There are various tokens of the internalist view. However, generally speaking, internalists are not always clear on what exact degree of similarity is required in order for conceptions to be the same.

and need not critically assess whether this theory of individuation is correct for reasons to be elaborated upon below. Obviously, it is important that conceptions co-refer for conceptions to not be different because if they do not share the same extension, then the conceptions are literally *about* disparate things. Moreover, on an internalist theory, if the conceptions in question do not have a certain amount of internal similarity, then we can be said to have different conceptions despite co-reference. However, being said to have the same conception does not necessarily mean that the constituents of our conceptions are exactly identical. For instance, my DOG conception may be constituted by BARKS, WAGS ITS TAIL, PLAYS FETCH, HAS FOUR LEGS, while your DOG conception is made up of BARKS, WAGS ITS TAIL, GUARDS THE HOUSE, and HAS FOUR LEGS. Even though we have a non-identical set of constituents to our conceptions, in this circumstance, internalists will still maintain that there is enough similarity such that we have conceptions that are not different. Here, one of us may potentially and legitimately claim that his or her conception of dogs is more accurate or correct than the other's. People can have non-identical conceptual constituents from each other, but as internalists claim, so long as there is a certain degree of similarity in their constituents and the conceptions co-refer, then they can be said to have the same conceptions.[57]

On the other hand, one may differentiate conceptions only by reference. Let us call this theory of individuation *the externalist theory*. On this view, what it means to

---

[57] For internalists and the notion of similarity, there must be a way of pinning down strict identity for the constituents, or else they need a non-constituent based notion of similarity that applies to the constituents.

have two conceptions that are not different is to have two conceptions that share the same reference, *tout court*. There is no additional requirement as with the internalist view to have a certain degree of similarity in cognitive content in order to be said to have the same conception. Moreover, what it means to have two differentiated conceptions on this theory is to have two conceptions that do not share the same reference, *tout court*. Once again, no additional requirement concerning cognitive content is relevant here. For example, one may be an externalist and maintain an informational theory of content. Recalling the informational view of content discussed in the first chapter, one has the concept *C* of the property *P* so long as one stands in the proper reliable causal relationship with *P*. So long as the tokening of both of our conceptions are reliably caused by *P*, then we have the same conceptions since they co-refer despite the fact that what constitutes our conceptions may be radically non-identical. As we can see, if the externalist view is correct and philosophers' and the folk's moral conceptions stand in the proper reliable causal relationship with the identical relevant property, then they all have conceptions that are not different, and the present potential objection at hand will not apply.

Now, as the present potential objection against the expert conceptual analysis of conceptions goes, if philosophers and the folk do indeed have different conceptions based on either an internalist or externalist theory, then their beliefs concerning the extension of the conceptions have different truth conditions. Since they have different truth conditions, like comparing apples and oranges, one cannot claim that professional

philosophers are better at the conceptual analysis of moral conceptions as compared to the folk when they analyze their own moral conceptions because the philosophers' moral conceptions and the folk's are not really the same such that such a comparison may be made. With different truth conditions, the philosopher and folk's respective intuitions that may at times be seemingly antithetical, both can be relativistically true. Due to this fact, there is no correct conception for the issue at hand towards which the two different conceptions may be compared. On the other hand, if they are the same conceptions based on either an internalist or externalist theory, then something needs to be said by the expert conceptual analyst of conceptions that can account for why they are the same. As we can see, the present entertained objection importantly relies on one's theory of conception individuation. While making this above potential objection against expert conceptual analysis of conceptions may be a possibility, in the first chapter, I was able to bracket the issue of conception individuation to the side and say that I will not provide a theory of conception individuation. Moreover, as we can see, the issue of conception individuation crucially at least relies on the notion of reference and what theory of reference one espouses. However, as also stated in the first chapter, I do not provide or espouse a theory of reference for moral conceptions in this dissertation, and I bracketed off any future serious discussion of reference or intentionality. Thus, for the above reasons, I do not pursue this potential objection against expert conceptual analysis of conceptions any further or any other idea that may include or rely upon propounding a theory of conception individuation and arguing for a theory of intentionality.

Now, one may wonder if not discussing conception individuation will mean that one cannot discuss certain important issues. For example, one cannot fully discuss interpersonal conception individuation, and one also will not be able to fully discuss intrapersonal conception individuation across time. For example, if what constitutes one of my conceptions at a particular moment changes diachronically over time, is it the case that I have the same conception that is not different across time?[58] The response to this is that this is exactly right that I will not be able to discuss such important issues. I do not discuss such conception individuation here, and therefore, do not discuss the above particular issues, but I do not foresee this to be problematic for the specified aims here. For, what I am abiding by here is a division of labor clarified by Machery in the first two chapters of his book *Doing Without Concepts.*[59] For mental representations, there is one project concerning what the nature of mental representations is. It is about what constitutes mental representations, or, in other words, it is about what the structure of mental representations is. This project can be thought of as falling under philosophical

---

[58] To note, recall that I discuss in the first chapter that for conception pluralism, Machery gives a kind of intrapersonal conception individuation theory for determining whether one may have several divergent conceptions of a single category at one particular moment in time. This view can be seen as a synchronic rather than diachronic theory. Machery's theory of individuation is not meant to be used as a way to determine whether intrapersonal synchronic co-referential bodies of knowledge can legitimately be compared to one another such that one body of knowledge may potentially be deemed to be more accurate than another. Rather, the purpose of his individuation theory is to determine how conceptions are stored in long term memory and recruited into working memory. Recall that such an individuation theory is not dependent on a theory of content but is dependent on what kinds of psychological processes underlie cognition. Therefore, there is no problem for Machery to discuss such individuation without having espoused a theory of content.

[59] Machery, ibid.

psychology or the interdisciplinary field of cognitive science. On the other hand, there is another project concerning mental representations that is about providing a theory of representation or content. This can be thought of as a project in the philosophy of mind.[60] Those primarily interested in the philosophy of science may pursue only the former project concerning mental representations, such as Machery, without providing a theory of content and the relevant theory of individuation. Moreover, one may pursue only the latter project, such as Fred Dretske, without providing an important theory on the nature of mental representations.[61] However, both projects need not be mutually exclusive, and one may grandly pursue both of them such as Prinz.[62] In this dissertation, I have specified that I will be pursuing only the former project of examining the nature of moral mental representations. I abide by this former aspect of the division of labor primarily for space concerns. For example, notice that it takes an entire book for Prinz to lay out his view on both projects for concrete conceptions in *Furnishing the Mind.* Then, since the nature of moral conceptions seems to be a more complex issue than concrete conceptions given that they may involve emotions, it actually takes Prinz three books to provide his overall theory on both projects for moral conceptions. As we do not have the space to pursue such a grand dual project here, I do not provide a theory of reference and

---

[60] This division of labor is analogous to the philosophy of emotions. One may think that the question of what constitutes an emotion is an area in philosophical psychology, while a theory of representation for emotions falls within the domain of the philosophy of mind.
[61] Fred Dretske, *Knowledge and the Flow of Information*, (Cambridge, MA: The MIT Press, 1981). Fred Dretske, "Misrepresentation," in *Belief: Form, Content and Function*, ed. By R. Bogdan, (Oxford: Oxford University Press, 1986), pp. 17-36.
[62] Prinz, *Furnishing the Mind*. Prinz, *Gut Reactions.* Prinz, *The Emotional Construction of Morals*.

conception individuation in this dissertation, and thus, I must leave aside for another time important issues that absolutely requires one to propound a theory of content. This will not be problematic so long as I do not discuss any issues or make any claims in this dissertation that absolutely require me to provide a theory of content. My general strategy of argument going forward is that when issues of conception individuation arise, I will discuss and cover the different major possibilities for conception individuation in the relevant circumstances without committing myself to the correctness of any particular possibility. I will argue that in each of the possibilities, philosophers still cannot credibly assume that they have justified their intuitions on the MJI/MJE or motivational Humeanism debates.

Another worry about the special authority of philosophers' intuitions is that we should expect to see some convergence in the intuitions of professional philosophers over time. Intuitions from conceptual analysts in thought experiments are taken to be justified beliefs. If our intuitions are arrived upon through a process that provides justification to our intuitions that they are true, then *ipso facto*, we should expect there to be convergence from philosophers over time. However, there appears to be no such convergence in the literature, especially when taking into account Bourget and Chalmers' survey. Therefore, at least a prima facie claim may be made against expert conceptual analysis in this

regards.[63] Even after hundreds upon hundreds of years and at times thousands upon thousands of years, a general unanimity is absent in the MJI/MJE and motivation Humeanism literature, and it appears no consensus is forthcoming. Some may respond that conceptual analysis is a very difficult task that takes time. Hence, philosophers need more time on the MJI/MJE and motivational Humeanism debates, and consensus may be forthcoming. However, even if conceptual analysis is a difficult task, the years that have already accrued are of such a striking magnitude that we have some justification to doubt that any consensus is forthcoming. Now, I do not understand this particular contention to be a definitive objection, but I believe it does raise the burden of proof against those who adhere to an experts-based view of conceptual analysis.

Now, one may wonder whether I am relying on a theory of conception individuation in this circumstance as I am discussing philosophers' interpersonal conceptions. Perhaps I now must be forced to provide a theory of content and conception individuation for moral conceptions. For, if philosophers have different conceptions either on an internalist or externalist account, then we cannot really say that there is any possibility in the first place for real convergence in their intuitions. First, if philosophers have different conceptions of, for example, AMORALIST, then there is no hope for expert conceptual analysis in the MJI/MJE debate because there cannot be any real genuine convergence in intuitions since philosophers technically have different conceptions that

---

[63] I say that a prima facie claim may be made because it is always possible that upon running experiments on philosophers' relevant conceptions, we may discover that there actually is convergence in intuitions on these two debates.

constitute their respective beliefs about amoralists; beliefs which have different truth conditions.[64] Presumably, beliefs that are really convergent must have the same truth conditions. Their beliefs cannot converge if the relevant conceptions that constitute their respective beliefs are different and cannot even be possibly compared to one another such that one's conception potentially may be said to be more accurate than the other's. Expert conceptual analysis at least implies convergence across philosophers and thus, requires philosophers to have relevant conceptions that are not different.

On the other hand, if philosophers' conceptions are the same and are not different, then this does not necessarily mean that there must be convergence in intuitions.[65] For, if there is no convergence in intuitions when philosophers' conceptions are the same and are not different, this simply means that what constitutes philosophers' conceptions do not match up in the appropriate way such that they have identical intuitions on a particular matter. What I will now attempt to show is simply that even though

---

[64] There is an ample amount of communication between people who have the same conceptions. However, in special cases where people have different conceptions, there may still be a smaller degree of communication between individuals with different conceptions given that there will most likely be some partial overlap in the constituents of their conceptions. There can be a degree of mutual intelligibility. However, there still can be no convergence since the intuitions have different truth conditions. To note, even if there is the possibility for convergence when two people have different conceptions, the metaethical literature in our two debates still appears to show that there is no such convergence.

[65] If one does not believe or think that it is intelligible that the relevant interpersonal conceptions within the MJI/MJE and motivational Humeanism debates may be different such that there can be no comparison between them where one conception can be said to be more accurate than another, then one maintains that the relevant conceptions are not different. However, even if this is the case, as we shall see, my prima facie objection about convergence (as well as any other relevant objections I make against conceptual analysis) still remains even if the conceptions are not different.

philosophers have the same relevant conceptions on either an internalist or externalist account, they can still have contrary intuitions. I will first demonstrate this by examining the scenario where philosophers' conceptions are the same based on an internalist theory.

Assuming that philosophers' conceptions are not different based on an internalist theory, the cognitive content of philosopher's conceptions have a certain degree of internal similarity with each other. However, this allows for the fact that there may or may not be any convergence in intuitions. For example, two philosophers' conceptions are not different and have a sufficient degree of similarity, but this still allows for the case that your conception AMORALIST is constituted by POSSIBLY COULD EXIST but my AMORALIST conception is made up of COULD NOT POSSIBLY EXIST.[66] Stated more simply, two philosophers' conceptions may be sufficiently similar to justify saying they are the same on an internalist account, but they differ with respect to the relevant beliefs about amoralists. We could have contrary intuitions on the MJI/MJE debate despite having AMORALIST conceptions that are not different because what relevantly constitutes our conceptions that drive our relevant intuitive judgments may not match up. Since our conceptions assumedly are not different, this allows for the possibility that one of our conceptions can be said to be more accurate than the other's. Assuming that philosophers' conceptions are not different, the metaethical literature does prima facie

---

[66] Requiring complete internal identity in order to have conceptions that are not different is an implausible view for an internalist because it is unlikely that many people will have complete internal identity. Thus, such a rigid theory of conception individuation will implausibly maintain that very few, if any, people will have the same conceptions that are not different.

support this point that there is no general convergence in intuitions. To note, if one claims that philosophers then have different conceptions of AMORALIST given the above constitution disparities that result in contrary intuitions, then we once again run into the aforementioned problem of philosophers having different conceptions such that there can be no convergence amongst philosophers. In summary of this paragraph, I have attempted to show that even if philosophers' conceptions are the same based on an internalist account, they can still have contrary intuitions, and there still can be no convergence.

For an externalist theory, let us assume that we both stand in the identical causal relation to the relevant property, but we may have some constituents of our AMORALIST conceptions in common. However, it may be the case that we do not have identical intuitions on the MJI/MJE debate because your conception is in part constituted by POSSIBLY COULD EXIST, but mine is in part constituted by COULD NOT POSSIBLY EXIST. Assuming that philosophers' AMORALIST conceptions are not different, the metaethical literature does prima facie support this point that there is no general convergence in intuitions. On the other hand, it could be the case that POSSIBLY COULD EXIST constitutes both of our conceptions while we both stand in the identical causal relation to the relevant property. As we can see, if philosophers' conceptions are not different on an externalist account, the issue of convergence in intuitions hinges upon the cognitive content of conceptions, and convergence does not solely hinge upon the (presently assumed) point that the conceptions are not different. In this paragraph, I have attempted to demonstrate

that even if philosophers' relevant conceptions are the same based on an externalist account, they can still have contrary intuitions, and there still can be no convergence.

As discussed, if philosophers' conceptions are different based on either an internalist or externalist theory, then there is no convergence in intuitions. Also, without having to adjudicate between whether an internalist or externalist theory is correct, if philosophers' conceptions are not different based on either an internalist or externalist account, then whether there is convergence in intuitions or not is dependent on what constitutes their conceptions and whether the constituents line up in the appropriate way. The metaethical literature appears to show that they do not. Hence, if philosophers' conceptions are not different, a prima facie objection may be made that there is still no convergence.

Now, while I attempt to avoid providing a theory of content and conception individuation, I am perfectly entitled to discuss what constitutes moral conceptions. This is what is of vital importance to my present objection against expert conceptual analysis, and it is what allows me to prima facie argue that there still appears to be no convergence in intuitions for philosophers even though we assume that their relevant conceptions are not different. Therefore, all in all, whether philosophers' conceptions are different or not, I can still raise my objection against expert conceptual analysis concerning the convergence in intuitions from philosophers in our two metaethical debates. Furthermore, I can do this while also avoiding the need to take a stand on a theory of

conception individuation since I argue that regardless of whether philosophers' conceptions are different or not, there still appears to be no convergence.

To note, I have provided and will soon further provide objections against conceptual analysis in this chapter, some objections of which have an interpersonal comparison of individual's conceptions built into them that may therefore seem to implicitly require me to provide a theory of conception individuation. However, the spirit of a number of such objections is that, like in my above objection, they crucially rely on what constitutes conceptions. What is important for such objections is the issue of what constitutes the conceptions, which is an issue I am perfectly entitled to speak about, and it is not and will not be required that I have a theory of conception individuation in hand in order to determine whether the relevant conceptions are different or not. For, regardless of whether the relevant conceptions are different or not, salient objections will still await the expert conceptual analyst on either one of these two horns of whether the relevant conceptions are different or not. For many of my objections, I take it that I can bypass the issue of conception individuation in like manner as compared to my above objection concerning the lack of convergence in philosophers' intuitions, *mutatis mutandis*. Furthermore, a similar route that may be taken in providing objections against conceptual analysis that still can bypass the issue of conception individuation is to first point out that if conceptions are different in a particular subject matter, there will be problems for the relevant conceptual analysis on the issue. Hence, for the sake of argument, we may grant the conceptual analyst that the relevant conceptions are the same

or not different. In granting this assumption, we still escape the need to argue for a theory of individuation, while then being able to object that there still are problems that await the conceptual analyst even if the conceptions are not different.

As another objection against expert conceptual analysis, there is the worry that intuition beliefs are products of one's cultural environment, upbringing by family, friends, and even heritable personality traits, where philosophers that have such divergent personal historical influences may resultantly have contradictory intuitions.[67] To note, this objection may also apply *mutatis mutandis* to the folk conceptual analysis of conceptions as well. Furthermore, for ease of linguistic expression, when discussing this particular objection, I presuppose that philosophers' conceptions are not different. However, the particular objection presented here also can be understood as saying that regardless of whether philosophers' conceptions are interpersonally different or not, their intuitions are still mere products of historical contingency, and this is epistemically problematic.

---

[67] Concerning conception individuation in this matter, expert conceptual analysis presupposes that there will be convergence in philosophers' intuitions. Therefore, it is imperative that philosophers' conceptions be the same. However, if philosophers conceptions are the same on either an internalist or externalist account, then what constitutes their respective conceptions do not seem to line up in the appropriate way such that they, generally speaking, do not appear to have convergence in intuitions. This apparent non-convergence can in part be causally explained by historical contingency. As will be stated shortly in the text, for ease of linguistic expression, when discussing my particular objection here, I presuppose that philosophers' conceptions are not different. However, the particular objection presented here can be understood as saying that regardless of whether philosophers' conceptions are interpersonally different or not, their intuitions may be mere products of historical contingency, and this is epistemically problematic.

Now, this objection is suggestive since it appears that philosophers do have apparently irreconcilable intuitions for our two metaethical subject matters. Moreover, it is further suggestive since we know from previous chapters that it is possible that many of our interpersonal moral conceptions that are the constituents of our a priori intuition judgments may themselves have constituents that are substantially non-identical.[68] Philosophers, like the folk, are part of the natural world, and setting aside indeterminism at the microphysical level, which if relevant will mean that our intuitions are generated by a random process, we live in a causal deterministic universe. Therefore, intuitions may be merely a product of historical contingency or historical accident. It may be the case that a priori intuitions are highly susceptible to being determined by epistemically pernicious and irrelevant factors without one even realizing it; factors that may be responsible for why two people with disparate histories have contradictory intuitions. If this is so, it at least provides a good partial explanation for the opposing intuitions moral philosophers apparently have, and it explains the apparent opposing conceptual constituents found in various philosophers for their various bodies of knowledge. This naturalistic-based objection has explanatory power. It is interesting to note that the

---

[68] While the conceptions experiments proving this were largely run on the folk, we may infer that these results also apply to moral philosophers on the MJI/MJE and motivational Humeanism debates given the apparently irreconcilable contrary intuitions philosophers hold in these debates and the Bourget and Chalmers study. Moreover, insofar as those who partake in the conceptual analysis of concepts have intuition beliefs in the head concerning the constituents of concepts, their mental representations of the concepts in a way are relevant to the present inquiry. Hence, conceptions conclusions concerning the fact that we may have non-identical bodies of knowledge or that we may have contrary conception constituents may still be relevant to the subject matter of the conceptual analysis of concepts.

famous poet and Oxford professor W. H. Auden, while observing philosophers at Oxford wrote:

> Oxbridge philosophers, to be cursory,
> Are products of a middle-class nursery:
> Their arguments are anent
> What Nanny really meant.[69]

In partial support of this overall objection, Eric Schulz et al. experimentally have shown that professional philosopher's intuitions on moral responsibility can be accurately predicted based on the possession of the heritable personality trait of extroversion.[70] Extroverts are more likely to be compatibilists.[71] Regardless of one's training, heritable traits that manifest themselves over a wide variety of environments and environmental influences, can be irrelevant causal factors to philosophers' intuitions. Now, while this experiment is not one that directly implicates our two metaethical issues at hand, it does provide empirical support for placing some level of doubt that intuitions for our two metaethical issues may also be products of some kind of historical contingency that directly questions the epistemic legitimacy of said intuitions. If certain intuitions on moral responsibility are based on some kind of historically contingency, then so may intuitions on the MJI/MJE and motivational Humeanism debates.

---

[69] W. H. Auden, *Academic Graffiti*, (New York: Random House), 1972, p. 25.
[70] E. Schulz, E. Cokely, and A. Feltz. "Persistent bias in expert judgments about free will and moral responsibility: A test of the expertise defense, *Consciousness and Cognition*, 20: (2011), pp. 1722-1731.
[71] A compatibilist is one who believes determinism is compatible with freedom.

Now, by no means do I understand this historical contingency objection to be a knock-down argument against the various forms of conceptual analysis that are currently under scrutiny. However, it does provide even a further burden of proof on those who believe that only expert's intuitions should be sought for conceptual analysis to provide empirical evidence to support their positive claim.

Some may object to this historical contingency argument that even scientific beliefs held by scientists are then in a sense historically contingent, but this does not necessarily impugn the justification of scientific beliefs. For example, if a biologist who believes in evolution was raised alternatively 700 years ago within the confines of the Christian doctrine before Darwin, then this person would believe in the antithetical view of divine creationism. Therefore, since a biologist's belief in evolutionary theory is subject to historical contingency – a contingency that potentially could have led to alternative contradictory beliefs if one's historical circumstances had been altered – but the biologist's belief in evolutionary theory still is justified, there also is no corresponding problem for ethicists. However, it seems that there is a disanalogy here between scientific or generally empirical-based beliefs generated from the scientific method, history, common observation such as in pre-modern psychology, the social sciences, and/or empirical testing versus those a priori intuitions coming from the armchair moral philosopher. Since they have different methodologies, where in our example, the evolutionary biologist justifies her claim in good part based on a wide body of empirical evidence rather than solely from a priori intuitions, the armchair ethicist

cannot claim that her intuition-based methodology is properly analogous to a generally empirical-based methodology.[72]

Fourth, there is a gathering consensus of experimental data on intuitions from professional philosophers that their intuitions are not immune from epistemically irrelevant and pernicious factors such as ordering effects and framing effects. If indeed philosophers' intuitions are more superior to the folk, then we should expect them to be less susceptible to such factors that influence intuitions.[73] For example, numerous philosophers such as Thomas Grundmann, Frank Hofmann, and Joachim Horvath have claimed that unlike the folk, a philosopher's intuition is protected from unconscious and

---

[72] I say 'in good part' in this sentence because there is also abstract theoretical work to be done in constructing the theory of evolution. Furthermore, an "empirical-based methodology" may also be a methodology that is used in those fields in which empirical evidence is difficult to find. Also, it may not be clear whether there is the kind of historical contingency in a priori disciplines such as in mathematics, where historical contingency is responsible for contradictory beliefs. However, even if math is based on historical contingency, it will be difficult for the armchair moral philosopher to make an argument by analogy with math to then say that expert metaethical intuitions are also justified despite contingency. For, there appears to be at least a prima facie disanalogy between math versus the MJI/MJE and motivational Humeanism debates. For, there appears to be generally strong consensus on mathematical issues from experts in math as instanced by the uniformity in various arithmetic, algebra, geometry, trigonometry, and calculus textbooks. However, there appears to be no such general consensus from experts when examining the literature of our two metaethical issues.

[73] Concerning conception individuation on this particular matter, expert conceptual analysts presuppose that philosophers and the folk have conceptions that are not different such that a proper comparison can be made between them on their skills of conceptual analysis. If they are different, then the expert conceptual analyst cannot really claim that their abilities at conceptual analysis on the relevant conception are superior to the folk. Moreover, as will be stated shortly, the experiments at hand still show that philosophers' intuitions are susceptible to cognitive biases, which places epistemic doubt on such intuitions. If they are not different, experiments suggest that the folk are less susceptible to cognitive biases than philosophers. For ease of linguistic expression, when discussing my particular objection here, I presuppose that philosophers' and folk's conceptions are not different.

unwanted biases.[74] They have made such a claim without providing the requisite experimental proof. Regardless, several studies have suggested that the above philosophers are wrong.[75] To note, for ease of linguistic expression, when discussing my particular objection here, I presuppose that philosophers' and folk's conceptions are not different. However, even if philosophers and the folk have different conceptions, the objection at hand still can apply, *mutatis mutandis*. For, epistemic doubt can still be raised against expert conceptual analysis because philosophers with their own different conceptions are still significantly subject to epistemically pernicious factors like ordering effects.

For example, in one study, Eric Schwitzgebel and Fiery Cushman have shown that philosophers' intuitions are susceptible to ordering effects.[76] Ordering effects are usually unconscious biases, where a participant's judgment on each item in a group of questions may potentially vary depending on what order the items are presented. At twenty-five major research universities, they ran a series of studies on 324 professional philosophers, 753 professional academic non-philosophers, and 1,389 non-academics

---

[74] T. Grundmann, "Some hope for intuitions: A reply to Weinberg," *Philosophical Psychology*, 23: (2010), pp. 481-509. F. Hofmann, "Intuitions, concepts, and imagination," *Philosophical Psychology*, 23: (2010), pp. 529-546. J. Hovarth, "How (not) to react to experimental philosophy," *Philosophical Psychology*, 23: (2010), pp. 447-480. See also: J. Wright, "On intuitional stability: The clear, the strong, and the paradigmatic," *Cognition* 115: (2010), pp. 491-503.

[75] Through verbal communication, Stephen Stich has said that he has several experimental philosophy papers in the works that demonstrate that philosophers are more susceptible to biases than the folk.

[76] Eric Schwitzgebel and Fiery Cushman, "Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers," *Mind & Language* 27: (2012), pp. 135-153.

with no graduate degree in any field. They provided participants with a list of moral scenarios for making moral judgments. For example, the first scenario was the lever case as discussed in the fourth chapter. Here, a train is about to run over five people, but you may divert the train onto a sidetrack by pulling a lever. Pulling the lever, however, will lead to the death of one person. The second case was also the previously discussed footbridge case. Recall that for this scenario, a train is about to run over five people, but you may push a large person over the footbridge that hovers over the track in order to stop the train and save the five. However, doing so will lead to this large person's death. Later on in the same study, as, for instance, question fourteen and fifteen, our experimenters re-asked participants both of these cases but reversed the order in which they were presented. In re-asking the questions, our experimenters slightly varied the cases, but kept the morally relevant features the same. For instance, they would use a runaway boat instead of a train that is about to run over five swimmers, but you can push a large person in the boat's path to stop the boat but end up killing the large person. The entire study was littered with several different kinds of pairs of cases like the lever and footbridge case, where the pairs of questions were re-asked in slightly altered form later in the same test, but in a different order. Upon running the tests, our experimenters ran the data to see whether or not the ordering of the scenarios played a substantial role in influencing and biasing the decisions made on the moral scenarios, where participants will give different answers on a type of question depending on what order it was

presented. They found that philosophers are more susceptible to ordering effects than non-philosophers:

> Both philosophers and non-philosophers showed significant order effects for all three types of scenario. In our summary measure of order effects across all scenario judgments, philosophers and ethics PhDs trended marginally higher than the comparison groups. Thus, philosophers showed no greater tendency than non-philosophers to use the consistent application of moral principles to reduce order effects on their scenario judgments.[77]

They discovered that ordering effects have a large scale influence on philosophers' intuitions on moral scenarios presumably without their awareness. Moreover, non-philosophers were impacted by the influence of ordering effects less than philosophers. Even though this study did not examine the intuitions from the MJI/MJE or motivational Humeanism debates, the red flag it raises for the relevant particular moral issues is still sufficient for us to have pause for concern on the general statement that intuitions from expert moral philosophers are more trustworthy than those from the folk. Hence, such a study is still relevant to our interests. For, it may be the case that there may be ordering effects for intuitions on our two moral psychology issues as well. Such a study provides empirical evidential support for this possible fact. It does create a level of doubt. However, since more studies need to be run on experts' intuitions to make a complete case, I do not take the studies to be conclusive proof that any experts-based view of conceptual analysis is fundamentally wrong. Nevertheless, they do lay down a further burden of proof against such philosophers. Such philosophers need to provide

---

[77] Schwitzgebel and Cushman, p. 148.

empirical evidence of such a strength that it supports their positive claim and overrides this and the previous burden of proof objections I have given. I take my objections to conjointly produce a substantial burden. Therefore, we may claim that the balance of the scales lies in favor of the fact that the conceptual analysis of concepts and expert conceptual analysis of conceptions appear to be undermined.[78] My above objections generally demonstrate that the use of such forms of conceptual analysis does not sufficiently justify a priori conclusions on the MJI/MJE and motivational Humeanism debates.

Once again, continuing to assume that philosophers and the folk do not have different conceptions, one may raise the objection that even if ordering effects influence professional philosophers more than the folk, when hypothetical scenarios are presented in isolation and not in any order with other hypotheticals, philosophers' intuitions are more reliable than the folk. However, this is a claim in which empirical evidence is directly relevant and must be presented in order to substantiate such a claim. Given the above Schwitzgebel and Cushman study, the burden of proof assuredly falls on the experts-based philosopher to demonstrate his or her point.

---

[78] While I have clarified that the conceptual analysis of concepts is to be considered as an experts-based view, if one then understands it to be a folk-based theory due to my objections against experts-based views, then there are further objections to this move. For, it is an empirical matter as to what the folk intuitions are as to the constituents of concepts. Relevant tests need to be conducted to see if there is any consensus of intuitions by the folk. Since no such studies have been run, we must remain agnostic. Furthermore, such folk intuitions may be grounded in historical contingency.

Given the above arguments against conceptual analysis, such as the historical contingency argument that applies to all forms of conceptual analysis, I have attempted to show that there are serious reservations against even pursing the a priori modal MJI/MJE and motivational Humeanism debates.

6.4 Conclusion

In the preceding five chapters of this dissertation, I have followed in the footsteps of Hume's *A Treatise of Human Nature*. Just as the three books that constitute the *Treatise* synchronically are of the human understanding, emotions, and morals, I likewise have examined conceptions or what Hume and Locke called 'ideas' that are the building blocks of thought and of how we understand the world. Moreover, I have then examined emotions and moral conceptions. While Hume draws from his first two books to conclude in his final book that moral conceptions are constituted by emotions, I similarly have drawn on the contemporary conceptions and emotions literature to provide an overall theory on the nature of moral conceptions that they can be constituted by prototypes, exemplars, theories, or emotions.

In this final chapter, we have explored any further philosophical implications the preceding chapters may have in ethics. In the beginning of this chapter, I have shown that the demise of the classical view has ramifications in normative ethical theory for those who presuppose that our moral mental representations can have classical structure. Moreover, I have contended that another implication is that there can be counterexamples

to people's conceptions. The existence of such counters does not necessarily mean that one's conception is false.

Finally, I have argued against those philosophers who use conceptual analysis in order to address the MJI/MJE and motivational Humeanism debates. By relying on such things as studies in previous chapters that establish the viability of the prototype, exemplar, theory, and emotion theories, I have contended that such philosophers engaged in the conceptual analysis of conceptions are not justified because they need to support their claims with experimental data on the conceptions that are relevant to the formal aspect. Finally, I have objected that those philosophers engaged in the expert analysis of concepts or conceptions are not sufficiently warranted in believing that philosophers' a priori intuitions on the MJI/MJE and motivational Humeanism debates are justified.

Bibliography

Adolphs, R. "Neural Systems for Recognizing Emotion." *Current Opinion Neurobiology* 12: 2002, 169-77.

Ahn, W., Kim, N., Lassaline, M., & Dennis, M. "Causal Status as a Determinant of Feature Centrality." *Cognitive Psychology*, 41: 2000, 361-416.

Anscombe, G.E.M. "Modern Moral Philosophy," *Philosophy*, 33: 1958, pp 1-19.

Aristotle. *Nichomachean Ethic.* , Ed. by Hugh Tredennik. Trans. by J.A.K. Thomson. New York: Penguin Books, 2003.

Armstrong, S.; Gleitman, L.; and Gleitman, H. "What Some Concepts Might Not Be." *Cognition*, 13, 1983.

Arsenio, W. "Children's conceptions of the situational affective consequences of sociomoral events." *Child Development* 59: 1988, 1611-22.

Auden, W. H. *Academic Graffiti*. New York: Random House, 1972, p. 25.

Audi, Robert. *The Good in the Right.* Princeton, NJ: Princeton University Press, 2004.

Aurelius, Marcus. *The Meditations of Marcus Aurelius Antoninus*. Trans. by A.S.L. Farguharsen.Oxford: Oxford University Press, 2008.

Ayer, A. J. 1936. *Language, Truth and Logic*. London: Gollancz.

Baird, J. and Astington, J. "The role of mental state understanding in the development of moral cognition and moral action." *New. Dir. Child Adolescent Development* 103: 2004, 37-49.

Baird, J. and Moses, L. "Do preschoolers appreciate that identical actions may be motivated by different intentions?" *Journal of Cognitive Development* 2: 2001, 413-448.

Barish, Kenneth. *Pride and Joy*. Oxford: Oxford University Press, 2012.

J. Baron and I.Ritov. "Intuitions about penalties and compensation in the context of tort law." *Journal of Risk and Uncertainty*, 7: 1993, 17-33.

Barrett, K., Zahn-Waxler, C., and Cole, P. "Avoiders Versus Amenders: Implication for the Investigation of Guilt and Shame During Toddlerhood?" *Cognition and Emotion* 7: 1993, 481-505.

Barsalou, Lawrence. "Ad hoc categories." *Memory and Cognition,* 1 (3): 1983, 211-227.

Barsalou, L. "On the indistinguishability of exemplar memory and abstraction in category representation." in *Advances in Social Cognition, Vol. III: Content and process Specificity in the Effects of Prior Experiences*. Ed. by T. Srull and R. Wyer, Jr. Hillsdale, NJ: Erlbaum, 1990, 61-88.

Barsalou, L., Solomon, K., Wu, L. "Abstraction in Perceptual Symbol Systems." *Philosophical Transactions of the Royal Society of London: Biological Sciences* 358 (2003): 1177-1187.

Barsalou, L., Simmons, W.K., Barbey, A., and Wilson, C.D. "Grounding conceptual knowledge in modality-specific systems." *Trends in Cognitive* Science. 7: 2003, 84-91.

Barsalou, L., Pecher, D., Zeelenber, R., Simmons, W.K., Hamann, S.B. "Multimodal simulation in conceptual processing." *Categorization inside and outside the lab: Essays in honor of Douglas L. Medin*, ed. W. Ahn, R. Goldstone, B. Love, A. Markman, Nad P. Wolff. Washington, D.C.: American Psychological Association, 2005.

Blackburn, Simon. "Errors and the Phenomenology of Value." *Morality & Objectivity: A Tribute to John Mackie*. London: Routledge & Kegan Paul, 1985.

Blackburn, Simon. *Spreading the Word: Groundings in the Philosophy of Language*. New York: Oxford University Press, 1984.

Blackburn, Simon. *Ruling Passions*. Oxford: Oxford University Press, 1998.

Blair, R. "A cognitive developmental approach to morality: Investigating the psychopath." *Cognition* 57 1995, 1-29.

Blair, R. "Moral Reasoning and the Child with Psychopathic Tendencies." *Personality and Individual Differences* 25, 1997, 731-9.

Borg, J. S. and Sinnott-Armstrong, W." Moral Judgments in Psychopaths," forthcoming.

Bourget, David and Chalmers, David. "What do Philosopher's Believe?" *Philosophical Studies*, forthcoming.

Boyd, Richard. "How To Be a Moral Realist." *Essays on Moral Realism*. Ed Geoffrey Sayre-Mccord. Ithaca, NY: Cornell University Press, 1988.

Boyd, Richard. "Kinds, Complexity and Multiple Realization," *Philosophical Studies* 95: 1999, pp. 67-98.

Boyd, Richard. "Realism, Anti-foundationalism, and the Enthusiasm for Natural Kinds," *Philosophical Studies* 61 (1-2): 1991, pp. 127-148.

Boyd, Richard. "What Realism Implies and What It Does Not," *Dialectica*43: 1989, pp. 5-29.

Brink, David. *Moral Realism and the Foundation of Ethics*. Cambridge: Cambridge University Press, 1989.

Brooks, L. "Nonanalytic Concept Formation and Memory for Instances."in *Cognition and Categorization*. Ed. by E. Rosch and B.B. Lloyd. Hillsdale, NJ: Lawrence Erlbaum associates, 1978.

Carey, Susan. *Conceptual Change In Childhood*. Cambridge, MA: The MIT Press, 1985.

Carey, Susan. "Knowledge Acquisition: Enrichment or Conceptual Change?" *The Epigenesis of Mind: Essays on Biology and Cognition*. Ed S. Carey and R. Gelman. Philadelphia, PA: Lawrence Erlbaum Associates, Inc., 1991.

Carey, Susan. *The Origin of Concepts*. Oxford: Oxford University Press, 2009.

Carlsmith, K., Darley, J., and Robinson, P. "Why do we punish? Deterrence and just deserts as motives for punishment." *Journal of Personality and Social Psychology* 83: 2002, 284-299.

Chi, M.T., Fletovich, P.J., and Glaser, R. "Categorization and representation of physics problems by experts and novices." *Cognitive Science* 5, 1981, 121-152.

Churchland, Paul. "The Neural Representation of the Social World." *Mind and Morals*. Ed. L. May, M. Friedman, A. Clark. Cambridge, MA: The MIT Press, 1996.

Churchland, Paul. *A Neurocomputational Perspective*. Cambridge, MA: The MIT Press, 1989.

Ciaramelli, E., Muccioli, M., Ladavas, E., and di Pellgrino, G. "Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex." *Social Cognitive and Affective Neuroscience* 2, 2007, 84-92.

Clark, Andy. "Connectionism, Moral Cognition, and Collaborative Problem Solving." *Mind and Morals*. Ed. L. May, M. Friedman, A. Clark. Cambridge, MA: The MIT Press, 1996.

Cleckley, H.M. *The Mask of Sanity: An Attempt to Reinterpret the So-Called Psychopathic Personality*, St. Louis, MO: The C.V. Mosby Company, 1941.

Colby, A., Kohlberg, L., Gibbs, J., and Lieberman, M. "A Longitudinal Study of Moral Judgment." *Monographs of the Society for research in Child Development* 48: 1983.

Coleman, Linda and Kay, Paul. "Prototype Semantics: The English Word Lie." *Language* 57:1 March 1981, 26-44.

Coley, J.,Medin, D., and Atran, S. "Does rank have its privilege? Inductive inferences within folk biological taxonomies." *Cognition* 64, 1997, 73-112.

Cushman, Fiery."Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment." *Cognition*, 108, 2008, 353-380.

Cushman, Fiery; Young, Liane; and Greene, Joshua. "Our Multi-system Moral Psychology: Towards a Consensus View." *The Oxford Handbook of Moral Psychology*. Ed J. Doris, G. Harman, S. Nichols, J. Prinz, W. Sinnott-Armstrong, S. Stich. Oxford: Oxford University Press, (forthcoming).

Damasio, Antonio. *Descartes' Error*. New York: Penguin, 1994.

Dancy, Jonathan. *Moral Reasons*. Oxford: Basil Blackwell, 1993.

Davidson, P.,Turiel, E., and Black, A. "The effect of stimulus familiarity on the use of criteria and justifications in children's social reasoning." *British Journal of Developmental Psychology* 1: 1983, 49-65.

Donagan, Alan. *The Theory of Morality*. Chicago: University of Chicago Press, 1979.

Dougherty, J.W.D. "Salience and relativity in classification." *American Ethnologist* 5: 1978, 66-80.

Doris, John. *Lack of Character*. Cambridge: Cambridge University Press, 2005.

Dretske, Fred. *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press, 1981.

Dretske, Fred. "Misrepresentation." *Belief: Form, Content and Function*. Ed. By R. Bogdan.Oxford: Oxford University Press, 1986, pp. 17-36.

Dwyer, Susan. "Moral Competence."*Philosophy and Linguistics*. Boulder, CO: Westview Press, 1999.

Ezcurdia, Maite. "The Concept-Conception Distinction." *Philosophical Issues*, Vol. 9, 1998.

Fehr, E. and Gachter, S. "Altruistic punishment in humans." *Nature* 415: 2002, 137-140.

Flanagan, Owen. "Ethics Naturalized: Ethics as Human Ecology." *Mind and Morals*. Ed. Larry May, Marilyn Friedman, and Andy Clark. Cambridge, MA: The MIT Press, 1996.

Flanagan, Owen. *Varieties of Moral Personality*. Cambridge, MA: Harvard University Press, 1991.

Fodor, Jerry. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press, 1998.

Fodor, Jerry. "The Current Status of the Innateness Controversy." in *Representations*. By J. Fodor. Cambridge, MA: The MIT Press, 1981.

Fodor, Jerry. "Having Concepts: a Brief Refutation of the Twentieth Century." *Mind and Language* 19: 2004, 29-47.

Fodor, Jerry. "Information and Representation." *Information, Language, and Cognition*. Ed P. Hanson. Oxford: Oxford University Press, 1991.

Fodor, Jerry. *The Language of Thought*. Cambridge, MA: Harvard University Press, 1975.

Fodor, Jerry. *LOT2*. Oxford: Oxford University Press, 2008.

Fodor, J., Garrett, M., Walker, E., and Parkes, C. "Against Definitions." *Cognition* 8: 1980, pp. 263-367.

Fodor, Jerry and Pylyshyn, Zenon. "Connectionism and Cognitive Architecture: A Critical Analysis." *Cognition* 28: 1988, 3-71.

Frege, G, "On Sense and Meaning." trans. M. Black.ed P. Geach and M. Black. *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Blackwell, 1953.

Frei, Jennifer and Shaver, Phillip. "Respect in close relationships: Prototype definition, self-report assessment, and initial correlates." *Personal Relationships* 9: 2002, 121-39.

Gallie, W.B. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society*. Vol.56, 1956, pp.167-198

Gibbard, Allan. *Thinking How to Live*. Cambridge, MA: Harvard University Press, 2003.

Gibbard, Allan. *Wise Choices, Apt Feelings*. Oxford: Clarendon Press, 1991.

Goldman, Alvin. "Ethics and Cognitive Science." *Ethics* Vol. 103; Jan 1993, 337-360.

Goodman, Nelson. "Seven Strictures on Similarity." In *Problems and Projects*. Indianapolis: Bobbs- Merrill, 1972.

Gopnik, Alison and Meltzoff, Andrew N.*Words, Thoughts, and Theories*. Cambridge, MA: The MIT Press, 1997.

Gorenstein, E. E. "Frontal Lobe Functions in Psychopaths." *Journal of Abnormal Psychology*, 91: 1982, p.368-379.

Greene, Joshua. "Dual-process Morality and the Personal/Impersonal Distinction: A Reply to McGuire, Langdon, Coltheart, and Mackenzie." *Journal of Experimental Social Psychology*, 2009.

Greene, Joshua. "From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?" *Nature Review Neuroscience* 4 2003: 847-50.

Greene, Joshua. "Reply to Mikhail and Timmons." *Moral Psychology Volume 3*.Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Greene, Joshua. "The Secret Joke of Kant's Soul." *Moral Psychology Volume 3*.Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Greene, Joshua. "Why are VMPFC Patients More Utilitarian?" *Trends in Cognitive Sciences* Vol. 11 No. 8, 2007.

Greene, Joshua; Morelli, Sylvia; Lowenberg, Kelly; Nystrom, Leigh; and Cohen, Jonathan. "Cognitive Load Selectively Interferes with Utilitarian Moral Judgment." *Cognition* 107, 2008, 1144-54.

Greene, J; Nystrom, L.E.; Engell, A.D. ;Darley, J.M., and Cohen, J.D. "The Neural Bases of Cognitive Conflict and Control in Moral Judgment," Neuron 44 (2004), 387-400.

Greene, J., Sommerville, R.B. Nystrom, L.E., Darley, J.M, and Cohen, J.D. "An fmri investigation of emotional engagement in moral judgment." *Science* 293, 2001, 2105-8.

Grundmann, T. "Some hope for intuitions: A reply to Weinberg." *Philosophical Psychology*, 23: 2010, pp. 481-509.

Haidt, Johnathan. "The emotional dog and its rational tail: A social intuitionist approach to moral judgment." *Psychological Review.108*, 814-834, 2001.

Haidt, Johnathan & Bjorklund, Fredrik. "Social Intuitionists Answer Six Questions about Moral Psychology." *Moral Psychology Volume 2*. Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Haidt, J., Koller, S., and Dias, M. "Affect, culture and morality, or is it wrong to eat your dog?" *Journal of Personality and Social Psychology* 65: 1993, 613-628.

Hales, S. D. *Relativism and the Foundations of Philosophy*. Cambridge, MA: MIT Press, 2006.

Hampton, James. "Inheritance of attributes in natural concept conjunctions." *Memory & Cognition*, 15: 1987, 55-71.

Hampton, James A. "An Investigation of the Nature of Abstract Concepts. "*Memory and Cognition* Vol.9 (2), 1981.

Hampton, James. "Polymorphous Concepts in Semantic Memory." *Journal of Verbal Learning and Verbal Behavior* 18: 1979, 441-61.

Hare, R. M. *The Language of Morals*. Oxford: Clarendon, 1952.

Hare, R. M. *Moral Thinking*. Oxford: Oxford University Press, 1982.

Harman, Gilbert. "Using a Linguistic Analogy to Study Morality." *Moral Psychology Volume 1*.Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Hauser, Marc D. *Moral Minds*. New York: Harper Collins Publishers, 2006.

Hauser,M., Cushman, F., Young, L., Kang-Xing Jin, R., and Mikhail, J. "A Dissociation Between Moral Judgments and Justifications." *Mind and Language*, Vol. 22 No. 1, February 2007, pp. 1-21.

Hauser, Marc D.; Young, Liane; & Cushman, Fiery."Reviving Rawls Linguistic Analogy: Operative Principles and the Causal Structure of Moral Actions." *Moral Psychology Volume 2*.Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Higginbotham, James. "Conceptual Competence."*Philosphical Issues*, Vol 9, 1998.

Hobbes, Thomas. *Leviathan*. Ed C. B. Macpherson. Harmondsworth: Penguin, 1981.

Hofmann, F. "Intuitions, concepts, and imagination." *Philosophical Psychology*, 23: 2010, pp. 529-546.

Horgan, Terrence and Timmons, Mark."Troubles for New Wave Moral Semantics: The 'Open Question Argument' Revived." *Philosophical Papers*, Vol. 21, No. 3 1992, 153-75.

Hovarth, J. "How (not) to react to experimental philosophy." *Philosophical Psychology*, 23: 2010, pp. 447-480.

Hull, Clark. "Quantitative aspects of the evolution of concepts." *Psychological Monographs, XXVIII*, 1920.

Hulme, C., Roodenrys, S., Brown, G., and Mercer, R. "The role of long-term memory mechanisms in memory span." *British Journal of Psychology* 86: 1995, pp. 527-536.

Hume, David. *An Enquiry Concerning the Principles of Morals*. Indianapolis, IN: Hackett Publishing Company, Inc., 1983.

Hume, David. *A Treatis of Human Nature*. Ed. By D. Norton and M. Norton. Oxford: Oxford University Press, 2000.

Hursthouse, Rosalind. *On Virtue Ethics*. Oxford: Oxford University Press, 1999.

Inhelder, Barbel and Piaget, Jean. *The Early Growth of Logic in the Child: Classification and Seriation*. London: Routledge and Kegan Paul, 1964.

Jackendoff, Ray. *Semantics and Cognition*. Cambridge, MA: The MIT Press, 1983.

Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly* 32, 1982, 127-136.

Jackson, Frank. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press, 1998, p. 31.

Johnson, C. and Keil, F. "Explanatory Understanding and Conceptual Combination." *Explanation and Cognition*, ed. F. Keil and R. Wilson. Cambridge, MA: The MIT Press, 2000, 327-59.

Johnson, Mark. *Moral Imagination*. Chicago: The University of Chicago Press, 1993.

Juslin, P., Jones, S., Olsson, H., and Winman, A. "Cue abstraction and exemplar memory in categorization." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29: 2003, 924-41.

Kahneman, D., Schkade, D., and Sunstein, C. "Shared outrage and erratic rewards: The psychology of punitive damages." *Journal of Risk and Uncertainty* 16: 1998, 49-86.

Kant, Immanuel. *The Critique of Pure Reason*. London: Macmillan, 1961.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Ed. by Mary Gregor. Cambridge: Cambridge University Press, 1997.

Keil, Frank C. *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: The MIT Press, 1989.

Keil, F., Smith, W., Simons, D, and Levin, D. "Two dogmas of conceptual empiricism: Implications for hybrid models of the structure of knowledge." *Cognition* 65: 1998, p. 103-135.

Kelly, D., Stich, S., Haley, K., Eng., S., Fessler, D. "Harm, Affect, and the Moral/Conventional Distinction." *Mind and Language*22: 2007, 117-131.

Kennett, Jeannett. "Empathy and Moral Agency." *The Philosophical Quarterly* 52: 2002, pp. 340-357.

Knobe, Joshua. "Intentional Action and Side Effects in Ordinary Language."*Analysis*.63: 2003, 190-193.

Knobe, Josua and Prasada, Sandeep. "Dual Character Concepts," unpublished data.

Kochanska, G., Gross, J., Lin, M., and Nichols, K. "Guilt in Young Children: Development, Determinants, and Relations With A Broader System of Standards. *Child Development* 73: 2002, 461-82.

Koechlin, E., Ody, C., and Kouneiher, F. "The Architecture of Cognitive Control in the Human Prefrontal Cortex." *Science* 302: 2003, 1181-5.

Koenigs, Michael and Tranel, Daniel." Irrational Economic Decision-Making after Ventromedial Prefrontal Damage: Evidence from the Ultimatum Game." *The Journal of Neuroscience* 27: 2007, 951-956.

Koenigs, M., Young, L., Adolphs, R.,Tranel, D., Cushman, F., Hauser, M., Damasio, A. "Damage to the prefrontal cortex increases utilitarian moral judgments." *Nature*, Vol. 446: 2007,908 - 911.

Kohlberg, Lawrence. *Essays on Moral Development, Vol. 1, The Philosophy of Moral Development*. New York: Harper & Row, 1981.

Komatsu, Lloyd. "Recent views of Conceptual Structure." *Psychological Bulletin* 112 (1992): 500-526.

Korsgaard, Christine. "Skepticism About Practical Reason." *The Journal of Philosophy*. 1986, p. 5-25.

Korsgaard, Christine M. *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.

Kovesi, Julius. *Moral Notions*. London: Routledge & Kegan Paul, 1967.

Kripke, Saul. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

Kruschke, John. "ALCOVE: An Exemplar-Based Connectionist Model of Category Learning." *Psychological Review* 99: 1992, 22-44.

Kunda, Z., Miller, D., and Claire, T. "Combining Social Concepts: The Role of Causal Reasoning. *Cognitive Science* 14 (1990): pp. 11-46.

Lakoff, George. *Women, Fire, and Dangerous Things.* Chicago: The Chicago University Press, 1987.

Lakoff, George and Nunez, Rafael. *Where Mathematics Comes From*. New York: Basic Books, 2001.

Laurence, Stephen and Margolis, Eric."Concepts and Cognitive Science." *Concepts*.Ed Eric Margolis and Stephen Laurence. Cambridge, MA: The MIT Press, 1999.

Lazarus, R. "On the primacy of cognition." *American Psychologist* 39: 1984, pp. 124-129.

Lazarus, R. and Alfert, E., "Short-circuiting of threat by experimentally altering cognitive appraisal." *Journal of Abnormal and Social Psychology* 69: 1964, 195-205.

Learner, Jennifer, Goldberg, Julie, and Tetlock, Philip. "Sober Second Thought." *Personality and Social Psychology Bulletin* 24, no. 6, 1998.

LeDoux, J. E. *The Emotional Brain*. New York: Simon and Schuster, 1996.

Lewis, Michael. "Emotional Competence and Development." *Improving Competence across the Lifespan. E*d. by D. Pushkar, W. M. Bukowski, A. E. Schwartzman, D.M. Stack and D. R. White. New York: Plenum, 1998, 27-36.

Lin, E. and Murphy, G. "The effects of background knowledge on object categorization and part detection." *Journal of Experimental Psychology* 50A: 1997, 25-48.

Locke, John. *An Essay Concerning Human Understanding*. Ed Kenneth P. Winkler. Indianapolis, IN: Hackett Publishing Company, Inc., 1996.

Lopez, A., Atran, S., Coley, J., Medin, D., and Smith, E. "The tree of life: Universal and cultural features of folkbiological taxonomies and inductions." *Cognitive Psychology* 32: 1997, 251-295.

Ludwig, K. "The epistemology of thought experiments: First person versus third person approaches." *Midwest Studies in Philosophy*, 31, 2007, pp. 128-159.

Machery, Edouard. "Concepts are not a Natural Kind." *Philosophy of Science* 72: (2005), 444-67.

Machery, Edouard. "Discovery and Confirmation in Evolutionary Psychology." *The Oxford Handbook of Philosophy of Psychology*. Ed. by J. Prinz. Oxford: Oxford University Press, forthcoming.

Machery, Edouard. *Doing Without Concepts*. Oxford: Oxford University Press, 2009.

Machery, Edouard. "How To Split Concepts: A Reply to Piccinini and Scott." *Philosophy of Science* 73, October 2006.

MacIntyre, Alasdaire. *After Virtue*, 2nd edition. Notre Dame, IN: Notre Dame Press, 1984.

Mackie, John. *Ethics: Inventing Right and Wrong*. New York: Penguin Books, 1991.

Maddock, R.J. "The Retrosplenial Cortex and Emotion: New Insights from Functional Neuroimaging of the Human Brain." *Trends Neuroscience* 22: 1999, 310-16.

Malt, B.C. "An online investigation of prototype and exemplar strategies in classification." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15: 1989, 539-555.

Martens, WHJ. "Antisocial and psychopathic personality disorders: causes, course and remission." *International Journal of Offender Therapy and Comparative Criminology*. 44(4): 2000, 406-430.

Martens, WHJ. *Psychopathy and Maturation*. PhD-thesis. Tilburg University. The Netherlands. Maastricht: Shaker Publishing, 1997.

McDowell, John. "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society*, suppl. Vol. 52: (1978).

McDowell, John. "Virtue and Reason," *Monist* 62: (1979).

McGinn, Colin. *Truth By Analysis*. Oxford: Oxford University Press, 2012.

McGuire, J., Langdon, R.,Coltheart, M., and Mackenzie, C."A reanalysis of the personal/impersonal distinction in moralpsychology research." *Journal of Experimental Social Psychology* 10, 2009.

McMahan, Jeff. "Moral Intuition." *Blackwell Guide to Ethical Theory*. Ed. by LaFollette, Hugh. Oxford: Blackwell, 2000.

Medin, D.,Altom, M., and Murphy, T.“ Given vs. induced category representations: Use of prototype and exemplar information in classification.” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10, 1984, 333-52.

Medin, Douglass and Schaffer, Marguerite. “A Context Theory of Classification Learning.” *Psychological Review* 85: 1978, 207-38.

Medin, D. and Schwanenfluegel, P. “Linear Separability in Classification Learning.” *Journal of Experimental Psychology: Human Learning and Memory* 7: 1981, 355-68.

Medin, D. and Shoben, E. “Context and structure in conceptual combination.” *Cognitive Psychology* 20: 1988, 158-90.

Medin, D., Lynch, E., and Solomon, K. “Are there kinds of concepts?” *Annual Review of Psychology* 51: (2000), 121-47.

Mendez, Mario. “What frontotemporal dementia reveals about the neurobiological basis of morality.” *Medical Hypothesis* 67: 2006, 411-418.

Mendez, M.F., Anderson, E., & Shapria, J.S. “An investigation of moral judgment in frontotemporal dementia.” *Cognitive and Behavioral Neurology*. 18, 4: 2005, 193-7.

Mercier, H. and Sperber, D. “Intuitive and reflective inferences.” *In two minds: Dual processes and beyond*. Ed. by Evans, J. & Frankish, K. Oxford: Oxford University Press, 2009.

Mikhail, John. “Moral Cognition and Computational Theory.” *Moral Psychology Volume 3*.Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Mikhail, John. “The Poverty of the Moral Stimulus.” *Moral Psychology Volume 1*.Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Mikhail, John. “Universal Moral Grammer: Theory, Evidence and the Future.” *Trends in Cognitive Sciences* Vol. 11 No. 4, 2007.

Mill, John Stuart. *Utilitarianism*.2nd edition.Edited by George Sher. Indianapolis, IN: Hackett Publishing Co., 2010.

Miller, David. “Justice and Boundaries.” *Politics, Philosophy & Economics* 8: 2009, pp. 291-309.

Miller, E.K. and Cohen, J.D. "An Integrative Theory of Prefrontal Cortex Function." *Annual Review Neuroscience* 24: 2001, 167-202.

Miller, G. A. "The Magical number seven, plus or minus two: Some limits on our capacity for processing Information." *Psychological Review* 63: 1956, pp. 81-97.

Moll, Jorge and De Oliveira-Souza, Ricardo." Moral Judgments, Emotions, and the Utilitarian Brain." *Trends in Cognitive Science* 11: 2007, 319-21.

Moll, Jorge and De Oliveira-Souza, Ricardo." Response to Greene: Moral Sentiments and Reason: Friends or Foes?" *Trends in Cognitive Science* 11: 2007, 323-4.

Moore, G. E. *Principia Ethica*, 5th edition. Cambridge: Cambridge University Press, 1968.

Murphy, Gregory. *The Big Book of Concepts*. Cambridge, MA: The MIT Press, 2002.

Murphy, Gregory. "Comprehending complex concepts." *Cognitive Science* 12, 1988, 529-62.

Murphy, Gregory. "Cue Validity and Levels of Categorization." *Psychological Bulletin* 91: 1982, 174-177.

Murphy, G. and Kaplan, A. "Feature distribution and background knowledge in category learning." *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology.* 53A: 2000, 962-982.

Murphy, G. and Medin, D." The Role of Theories in Conceptual Coherence." *Psychological Review* 92, 1985.

Nado, Jennifer, Kelly, Daniel, Stich, Stephen. "Moral Judgment." *Routledge Companion to the Philosophy of Psychology*. Ed. by John Symons & Paco Calvo. New York: Routledge, 2009.

Nagel, Thomas. *The Possibility of Altruism*. Oxford: Clarendon Press, 1970.

Nichols, Shaun. "After Objectivity: An Empirical Study of Moral Judgment." *Philosophical Psychology* 17: 2004, 5-28.

Nichols, Shaun. "Innateness and Moral Psychology." *The Innate Mind*. Ed Peter Carruthers, Stephen Laurence, and Stephen Stich. Oxford: Oxford University Press, 2006.

Nichols, Shaun. "Norms with feeling: Toward a psychological account of moral judgment." *Cognition*84: 2002, 223-236.

Nichols, Shaun. "Sentimentalism Naturalized." *Moral Psychology, Vol. 2*. Ed. by Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008, p. 255-274.

Nichols, Shaun. *Sentimental Rules*. Oxford: Oxford University Press,2004.

Nisan , M . "Moral norms and social conventions: A cross-cultural comparison." *Developmental Psychology* , 23: 1987, 719 – 725 .

Nosofsky, R. "Exemplars, prototypes and similarity rules." in *From Learning Theory to connectionist Theory: Essays in Honor of W. K.* Estes. Ed. by A. Healy, S. Kosslyn, and R. Shiffrin. Hillsdale, NJ: Erlbaum, 1992, p. 149-68.

Nucci, L. *Education in the Moral Domain*. Cambridge: Cambridge University Press, 2001.

Nucci, L. and Nucci, M. "Children's Social Interactions in the Context of Moral and Conventional Transgressions." *Child Development*, 53: 1982, 403-12.

Nucci, Larry and Turiel, Elliot. "Social Interactions and the Development of Social Concepts in Preschool Children." *Child Development*, 49: 1978, 400-407.

Nucci, Larry, Turiel, Elliot, and Encarnacion-Gawrych, Gloria. "Children's Social Interactions and Social Concepts." *Journal of Cross-Cultural Psychology*. 14: 1983, pp. 469-487.

Nussbaum, Martha. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press, 2001, 4.

Osherson, D., and Smith, E."On the Adequacy of Prototype Theory as a Theory of Concepts." *Cognition*9, 1981.

Osherson, D.N., Smith, E.E., Wilkie, O., Lopez, A., and Shafir, E. "Category-based induction." *Psychological Review* 97, 1990, 185-200.

Paharia, N., Kassam, K., Greene, J., Bazerman, M. "Dirty work, clean hands: The moral psychology of indirect agency." Organizational Behavior and Human Decision Processes, 2009.

Park, John. "Prototypes, Exemplars, and Theoretical & Applied Ethics," *Neuroethics*, forthcoming.

Park, John. "Theories of Moral Concepts," *Truth Matters,* ed. by Lambert Zuidervaart, forthcoming.

Peacocke, Christopher. *A Study of Concepts*. Cambridge, MA: The MIT Press, 1992.

Phan, K.L, Wager, T., Taylor, S.F., and Liberzon, I. "Functional Neuroanatomy of Emotion: A meta- analysis of Emotion Activation Studies in Pet and fmri." *Neuroimage*16: 2002, 331-348.

Piaget, Jean. *The Moral Judgment of the Child*. New York: Free Press, 1932.

Piccinini, Gualtiero and Scott, Sam." Splitting Concepts." *Philosophy of Science* 73: (2006), 390-409.

Pillutla, M., Murnighan, J. "Unfairness, anger, and spite: emotional rejections of ultimatum offers." *Org. Behav. Hum. Dec. Proc* 68: 1996, 208-224.

Plato. *Protagoras*. In *Plato Complete Works*. Ed. by John M. Cooper. Indianapolis: Hackett Publishing Company, 1997, 358c.

Plato. *The Republic*, 2nd edition. Trans. by Lee, Desmond. London: Penguin Books, 1974, pp. 18-19.

Posner, M. and Keele, S. "On the Genesis of Abstract Ideas." *Journal of Experimental Psychology* 77: 1968, 353-363.

Prinz, Jesse. *The Emotional Construction of Morals*. Oxford: Oxford University Press, 2008.

Prinz, Jesse. *Furnishing the Mind*. Cambridge, MA: The MIT Press, 2002.

Prinz, Jesse. *Gut Reactions*. Oxford: Oxford University Press, 2004.

Prinz, Jesse. "Resisting the Linguistic Analogy: A Commentary on Hauser, Young, and Cushman." *Moral Psychology Volume 2*. Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Prinz, Jesse. "The return of concept empiricism." *Categorization and Cognitive Science*, ed. By Cohen, H. and Leferbvre, C. Cambridge, MA: Elsevier Science, 679-94.

Prinz, Jesse and Clark, Andy. "Putting Concepts to Work: Some Thoughts for the Twentyfirst Century," *Mind & Language* 19: 2004, 57-69.

Proffitt, J., Coley, J. and Medin, D. "Expertise and category-based induction." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26, 2000, 811-28.

Putnam, Hillary. "The Meaning of "Meaning."" *Philosophical Papers, Vol 2: Mind, Language, and Reality*. By H. Putnam. Cambridge: Cambridge University Press, 1975.

Quine, W.V.O. "Natural Kinds." *Naming, Necessity, and Natural Kinds*. Ed. S.P. Schwarz. Ithaca, NY: Cornell University Press, 1977, 155-75.

Quine, W.V.O. "Two Dogmas of Empiricism," *The Philosophical Review* 60: 20-43, 1951.

Railton, Peter. "Moral Realism," *Philosophical Review*, 95, 1986

Railton, Peter. "Naturalism and Prescriptivity." *Social Philosophy and Policy* 7, 1989.

Ramnani, N. and Owen, A.M. "Anterior Prefrontal Cortex: Insights Into Function From Anatomy and Neuroimaging." *National Review Neuroscience* 5: 2004, 184-94.

Rawls, John. *A Theory of Justice*, 2nd edition. Cambridge, MA: Belknap Press, 1971.

Rehder, B. "When causality and similarity compete in category-based property induction." *Memory & Cognition* 34: 2006, 3-16.

Rey, Georges. "Concepts and Stereotypes." *Cognition* 19, 1983.

Rips, Lance. "Inductive judgments about natural categories." *Journal of Verbal Learning and Verbal Behavior*. 14. 1975, 665-81.

Rips, L., Shoben, E., and Smith, E." Semantic distance and the verification of semantic relations." *Journal of Verbal Learning and Verbal Behavior* 12: (1973), p. 1-20.

Rosch, Eleanor. "Cognitive representations of semantic categories." *Journal of Experimental Psychology: General.* 104: 1975, 192-233.

Rosch, Eleanor. "Natural Categories." *Cognitive Psychology* 4: 1973, 328-350.

Rosch, Eleanor. "Principles of Categorization." *Cognition and Categorization*. Ed Eleanor Rosch and B. Lloyd. Philadelphia, PA: Lawrence Erlbaum Associates, Inc., 1978.

Rosch, Eleanor and Mervis, Caroline. "Family resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 7: 1975, 573-605.

Rosenberg, Alex. *Philosophy of Science, 2nd Edition*. New York: Routledge, 2005.

Rosch, E.,Mervis, C., Gray, W., Johnson, D., and Boyes-Braem, P. "Basic Objects in Natural Categories." *Cognitive Psychology* 8: 1976, 382-439.

Roskies, Adina. "Are Ethical Judgments Intrinsically Motivational? Lesson from 'Acquired Sociopathy.'" *Philosophical Psychology* 16: 2003, pp. 51-66.

Roskies and Sinnott-Armstrong, W. "Between a Rock and a Hard Place: Thinking about Morality." *Scientific American Mind* 19: 2008.

Ross, L., Lepper, M., Hubbard, M. "Perseverance in self perception and social perception: biased attributional processes in the debriefing paradigm." *Journal of Personality and Social Psychology* 32: 1975, 880-892.

Sanfey, A.,Rillig, J., Aronson, J., Nystrom, L, and Cohen, J. "The neural basis of economic decision-making in the ultimatum game." *Science* 300: 2003, 1755-1758.

Scanlon, Thomas. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

Scherer, K. "Studying the emotion-antecedent appraisal process: An expert system approach." *Cognition and Emotion* 7: 1993, 325-356.

Schnall, Simone, Benton, Jennifer, and Harvey, Sophie. "With a Clean Conscience: Cleanliness Reduces the Severity of Moral Judgments." *Psychological Science*, 19, 2008, 1219-1222.

Schnall, Simone; Haidt, Jonathan; Clore, Gerald; and Jordan, Alexander. "Disgust as Embodied Moral Judgment." *Personality and Social Psychology Bulletin*. 34: 2008, 1096-1109.

Schroeder, T., Roskies, A., Nichols, S. "Moral Motivation." *The Moral Psychology Handbook*. Ed. by J. Dorris. Oxford: Oxford University Press, 2010.

Schulz, E., Cokely, E., and Feltz, A. "Persistent bias in expert judgments about free will and moral responsibility: A test of the expertise defense." *Consciousness and Cognition*, 20: 2011, pp. 1722-1731.

Schwarz, Norbert, and Clore, Gerald. "Mood, Misattribution, and Judgments of Well-Being." *Journal of Personality and Social Psychology* 45 no. 3, 1983.

Shafer-Landau, Russ. *Moral Realism: A Defense*. Oxford: Oxford University Press, 2003.

Shweder, R.,Mahapatra, M., and Miller, J. "Culture and Moral Development." *The Emergence of Morality in Young Children*. Ed. by Kagan, J. and Lamb, S. Chicago: University of Chicago Press, 1987.

Schwitzgebel, Eric and Cushman, Fiery. "Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers." *Mind & Language* 27: 2012, pp. 135-153.

Siebert, M., Markowitsch, H., Bartel, P. "Amygdala, affect and cognition: evidence from 10 patients with Urbach–Wiethe disease." *Brain* 126: 2003, 2627-2637

Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs*, 1972: 229-243.

Sinnott-Armstrong, Walter. "Framing Moral Intuitions." *Moral Psychology Vol. 2*. Ed. by W. Sinnott-Armstrong, Cambridge, MA: The MIT Press, 2008.

Sinnott-Armstrong, Walter. *Moral Dilemmas*. Oxford; New York, NY: Blackwell Publishing, 1988.

Sloman, Steven. "Feature-based induction." *Cognitive Psychology* 25: 1993, 231-280.

Sloman, Steven. "When explanations compete: The role of explanatory coherence on judgments of likelihood." *Cognition* 52, 1994, 1-21.

Smetana, Judith. "Preschool Children's Conceptions of Moral and Social Rules." *Child Development*. Vol. 52, No. 4: Dec. 1981, pp. 1333-1336.

Smetana, J. "Understanding of social rules. "*The Development of Social cognition: The Child as Psychologist*. Ed. by M. Bennett. New York: Guilford Press, 1993.

Smith, C. A. and Lazarus, R. S. "Appraisal components, core relational themes, and the emotions." *Cognition and Emotion* 7: 1993, pp. 233-269.

Smith, Edward and Medin, Douglas. *Categories and Concepts*. Cambridge, MA: Harvard University Press, 1981.

Smith, Edward and Osherson, Daniel. "Conceptual Combination with Prototype Concepts." *Cognitive Science* 8 (1984): 337-61.

Smith, E., Osherson, D., Rips, L., and Keane, M. "Combining prototypes: A selective modification model," *Cognitive Science* 12: (1988), 485-527.

Smith, E.,Shoben, E., and Rips, L. "Structure and Process in Semantic Memory: A Featural Model for Semantic Decisions." *Psychological Review* 81: 1974, 214-241.

Smith, J.D. and Minda, J.P. "Prototypes in the mist: The early epochs of category learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: 1998, 1411-36.

Smith, J.D., Murray, M.J., and Minda, J.P. "Straight talk about linear separability." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23: 1997, 659-680.

Smith, Kyle, Smith, Seyda, and Christopher, John. "What Defines the Good Person?" *Journal of Cross-Cultural Psychology* 38: 2007, 333-360.

Smith, Michael. *The Moral Problem*. Oxford: Blackwell Publishers, 1994.

Snarey, J. "Cross-Cultural Universality of Social-Moral Development: A Critical Review of Kohlbergian Research." *Psychological Bulletin*, 97: 1985, 202-32.

Solomon, K., Medin, D., Lynch, E. "Concepts Do More Than Categorize." *Trends in Cognitive Sciences* 3 (1999): 99-105.

Sripada, Chandra Sekhar. "Nativism and Moral Psychology." *Moral Psychology Volume 1*. Ed Walter Sinnott-Armstrong. Cambridge, MA: The MIT Press, 2008.

Stevenson, Charles. *Ethics and Language*. New Haven and London: Yale University Press, 1944.

Stich, Stephen. "Moral Philosophy and Mental Representation." *The Origin of Values*. Ed. by Hechter, M.,Nadel, L., and Michod, R. New York: Aldine de Gruyer, 1993, pp. 215-228.

Strichartz, Abigail and Burton, Roger. "Lies and Truth: A Study of the Development of the Concept." *Child Development*.61:1 Feb., 1990, 211-20.

Sweetser, Eve. "The definition of lie: An examination of the folk models underlying a semantic prototype." In *Cultural Models in Language and Thought*, ed. D. Holland and N. Quinn Cambridge: Cambridge University Press, 1989, 43-66.

Tanaka, J.W. and Taylor, M.E. "Object categories and expertise: Is the basic level in the eye of the beholder?" *Cognitive Psychology* 15, 1991, 121-149.

Thomson, Judith Jarvis. "A Defense of Abortion." *Philosophy and Public Affairs* 1:1 Autumn 1971: 47-66.

Turiel, E. *The Development of Social Knowledge*. Cambridge: Cambridge University Press, 1983.

Turiel, E."Distinct conceptual and developmental domains: Social convention and morality." *Nebraska Symposium on Motivation*. Ed. by H. Howe and C. Keasey. Lincoln: University of Nebraska Press, 1979.

Turiel, E., Killen, M., and Helwig, C."Morality: It's structure, functions, and vagaries." *The Emergence of Morality in Young Children*. Chicago: The University of Chicago Press, 1987.

Tversky, Amos. "Features of Similarity." *Psychological Review* 84, 1977, 327-350.

Tversky, A. &Kahneman, D. "The framing of decisions and the psychology of choice." *Science* 185: 1981, 453-8.

Valdesolo, P.&DeSteno, D. "Manipulations of emotional context shape moral judgment," *Psychological Science*17, 2006, 476–477.

Van Wout, M., Kahn, R.,Sanfey, A.,and Aleman, A. "Affective state and decision-making in the ultimatum game." *Experimental Brain Resonance* 169: 2006, 564-568.

Walker, Lawrence and Hennig, Karl."Differing Conceptions of Moral Exemplarity: Just, Brave, and Caring." *Journal of Personality and Social Psychology* 86: 2004, 629-647.

Walker, Lawrence and Pitts, Russell."Naturalistic Conceptions of Moral Maturity." *Developmental Psychology* 34: 1998, 403-418.

Ed. by Wallace, G. and Walker, A. *The Definition of Morality*, (London: Methuen & Co., 1970).

Wattenmaker, W.D., Dewey, G.I., Murphy, T.D., and Medin, D.L. "Linear separability and concept learning: Context, relational properties, and concept naturalness." *Cognitive Psychology* 18, 1986, 158-94.

Wheatley, Thalia and Haidt, Jonathan. "Hypnotic Disgust Makes Moral Judgments More Severe." *Psychological Science* 16:10: 2005, 780-784.

Weiskopf, Daniel. "Atomism, Pluralism, and Conceptual Content." *Philosophy and Phenomenological Research*, 79: 2009, pp. 130-162.

Weiskopf, Daniel. "Concept empiricism and the vehicles of thought." *Journal of Consciousness Studies* 14: (2007), 156–183.

Weiskopf, Daniel. "The Plurality of Concepts."*Synthese* 169: (2009), 145-173.

Whittlesea, B.W.A., Brooks, L.R., and Westcott, C. "After the learning is over: Factors controlling the selective application of general and particular knowledge." *Journal of Experimental Psychology: Learning, Memory, and* Cognition 20: 1994, 259-74.

Williams, Bernard. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press, 1985.

Williams, Bernard. "Moral Luck." *Moral Luck*. Cambridge: Cambridge University Press, 1981, pp. 22-39.

Williams, Bernard ."Persons, Character, and Morality." *Moral Luck.* Cambridge, Cambridge University Press, 1981, pp. 1-19.

Williamson, T. "Armchair philosophy, metaphysical modality and counterfactual thinking." *Proceedings of the Aristotelian Society*, 105, 2005, p. 1-23.

Wisniewski, E." When concepts combine." *Psychonomic Bulletin & Review* 4, 1997, 167-83.

Wisniewski, E. and Love, B. "Relations versus properties in conceptual combination." *Journal of Memory and Language* 38, 1998, 177-202.

Wittgenstein. *Philosophical Investigations*. Trans by G. Anscombe. New York: Prentice-Hall, Inc., 1953.

Wong, David B. *Natural Moralities*. Oxford: Oxford University Press, 2006.

Wright, J. "On intuitional stability: The clear, the strong, and the paradigmatic." *Cognition* 115: 2010, pp. 491-503.

Yamauchi, T. and Markman, A.B." Category learning by inference and classification." *Journal of Memory and Language* 39: 1998, 124-48.

Young, L., Camprodon, J., Hauser, M.,Pascual-Leone, A., and Saxe, R. "Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments." *PNAS*, 2010.

Young, L., Cushman, F., Hauser, M., and Saxe, R. "The neural basis of the interaction between theory of mind and moral judgment." *Proceedings of the National Academy of Sciences of the United States of America*, 104, 2007, 8235-8240.

Young, Liane and Saxe, Rebecca. "The neural basis of belief encoding and integration in moral judgment." *NeuroImage*, 2008.

Zajonc, Robert. "On the Primacy of Affect." *American Psychologist* 39: 1984, p. 117-123.

Biography

John Jung Park was born in Indianapolis, Indiana on October 2nd, 1978. He received a BA in Philosophy at DePauw University in the Spring of 2001, an MA in Philosophy at Western Michigan University in the Spring of 2005, and is a PhD candidate in Philosophy at Duke University.

Titles of publications: "Folk Moral Relativism," "Prototypes, Exemplars, and Theoretical & Applied Ethics," "The Hard Problem of Consciousness & the Progressivism of Scientific Explanation," "Theories of Concepts & Moral Truth."

A list of scholarships and awards:

- Graduate Student Essay Prize, Appalachian Regional Student Philosophy Colloquium, East Tennessee State University
- Anne T. and Robert M. Bass Fellowship, Duke University
- Franklin Humanities Institute Seed Grant, Duke University.
- Summer Research Fellowship, Duke University.
- Graduate Student Essay Prize, Truth Matters Conference, University of Toronto.
- James B. Duke Fellowship, Duke University.
- University Graduate Research & Creative Scholar Award, Western Michigan University
- Heraclitean Society Essay Contest, First Prize, Western Michigan University.

- Graduate Assistantship, Center of the Study of Ethics in Society at Western Michigan University.

- Research Assistantship, Center of the Study of Ethics in Society at Western Michigan University.